MARTA SKORKO

Fizyka

podręcznik dla studentów wyższych technicznych studiów zawodowych dla pracujących

Wydanie VI



WARSZAWA 1979
PAŃSTWOWE WYDAWNICTWO NAUKOWE

Okładkę projektowała
Maria Łuszczyńska

ISBN 83-01-00280-8

Spis rzeczy

CZĘŚĆ V. WYBRANE ZAGADNIENIA Z FIZYKI WSPÓŁCZESNEJ

Pamięci Męża Mojego
Współautora skryptowej wersji
niniejszego podręcznika
poświęcam

Przedmowa

Na pierwszym wykładzie zwracam się zwykle do studentów nowych roczników z życzeniem powodzenia w pracy nad fizyką. Niech mi wolno będzie takie samo życzenie skierować do Czytelników niniejszego podręcznika — słuchaczy wyższych technicznych studiów zawodowych dla pracujących.

Nie będę ukrywała, że czeka ich ciężka praca, wymagająca dużego wysiłku. Niewątpliwie pocieszające dla nich jest to, że zdaję sobie sprawę z trudności, jakie przed nimi stoją. Pamiętam m.in. o tym, że książka jest przeznaczona dla studentów, którzy wskutek kilkuletniej przerwy w nauce wypracowują sobie od nowa technikę uczenia się, że ich zasób wiedzy, oparty na wiadomościach wyniesionych ze szkoły średniej, jest znacznie mniejszy niż zasób wiedzy słuchaczy studiów dziennych, że łączą oni studia z pracą zawodową, co pociąga za sobą szereg oczywistych konsekwencji.

Te względy zdecydowały w dużej mierze o charakterze podręcznika. Wybór zagadnień i układ materiału były narzucone obowiązującym programem. Rozwinięcie zaś poszczególnych zagadnień, zgodnie z moimi zamierzeniami, miało zapewnić zrozumiałość treści, wdrażać do logicznego rozumowania jak również wyrabiać szerszy pogląd na świat i zjawiska w nim zachodzące.

Ze względu na to, że kurs fizyki na studiach zawodowych rozpoczyna się w semestrze drugim, możliwe jest operowanie w podręczniku elementami matematyki wyższej stopniowo w coraz szerszym zakresie. Aby uczynić tekst bardziej zrozumiałym, w dowodach matematycznych nie są pomijane przejściowe ogniwa rozumowania.

Do ważnych elementów podręcznika należą pytania i zadania. Pytania służą przede wszystkim do kontroli zdobytych wiadomości, jak również pomagają w wypracowaniu własnej techniki uczenia się przez podsuwanie różnych metod przyswajania wiadomości. Liczne zadania, wprowadzone do tekstu i zebrane na końcu każdego rozdziału dobrane są w ten sposób, aby były ilustracją przerobionego materiału, pomocą w jego opanowaniu, a zarazem aby pozwalały kontrolować stopień jego zrozumienia i przyswojenia. Zadania są typowe, takie, jakie można znaleźć w zbiorach zadań z fizyki przeznaczonych dla studiów technicznych.

Również jako pomoc w nauce pomyślane są wszystkie wprowadzone do podręcznika „Zestawienia wiadomości", „Tabele porównawcze", „Podsumowania" itp. Należy je traktować jako przykłady różnych sposobów syntetycznego ujmowania trudniejszych zagadnień.

Ze względów wymienionych w pierwszej części przedmowy nie brak też w podręczniku zwartego przedstawienia najistotniejszych wiadomości ze szkoły średniej.

W kilku przypadkach rozwinięte zostały szerzej uwagi dotyczące historii rozwoju fizyki, potraktowane jako ilustracja rozwoju myśli badawczej.

W podręczniku zastosowano: symbole wielkości fizycznych zgodne z zaleceniami Międzynarodowej Unii Fizyki Czystej i Stosowanej*, jednostki układu SI i jednostki przejściowo dopuszczone do stosowania oraz nazwy wielkości i skróty jednostek zgodne z podanymi w aktach prawnych obowiązujących w kraju w okresie przygotowywania podręcznika do druku**. Zestawienie używanych symboli i tablice przeliczeniowe jednostek podane zostały na końcu książki.

Na zakończenie niech mi wolno będzie wyrazić podziękowanie Kolegom z Katedry Fizyki Wydziału Chemicznego Politechniki Łódzkiej, dr Halinie Kurczewskiej i mgrowi Józefowi Świątkowi, którzy przekazali mi swoje uwagi dotyczące kilku rozdziałów. Pani Zofii Kalinowskiej dziękuję za bardzo staranne przygotowanie maszynopisu książki.

Kończę apelem do Czytelników, aby nasuwające się im wątpliwości, uwagi krytyczne itp. zechcieli przekazać na moje imię pod adresem Wydawnictwa

MARTA SKORKO

---

* *Symbols, Units and Nomenclature in Physics*, opr. przez International Union of Pure and Applied Physics, S.U.N. Commission, wyd. Unesco, 1965.

** Monitor Polski, Nr 74/66.

Wiadomości wstępne

1. Przedmiot fizyki i metody badań

Fizyka należy do nauk przyrodniczych. Przedmiotem jej badań są zjawiska zachodzące w przyrodzie martwej, celem badań — poznanie w sposób możliwie najbardziej wszechstronny praw rządzących tymi zjawiskami.

Wśród podstawowych bodźców, które przyczyniły się do rozwoju fizyki, wymienić należy przemożny instynkt poznawania otaczającego świata, jak również — a może nawet przede wszystkim — praktyczne potrzeby ludzi.

Podręcznik niniejszy przeznaczony jest dla przyszłych inżynierów. Warto więc od razu we wstępie podkreślić ścisły związek fizyki z techniką. Można tu mówić o swoistym sprzężeniu zwrotnym: potrzeby techniki prowadzą do odkryć fizycznych, a te z kolei niejednokrotnie umożliwiają dalsze osiągnięcia techniczne. Czasem odkrycia fizyczne o charakterze na pozór czysto naukowym zapoczątkowują szeroki rozwój pewnych dziedzin techniki. Wystarczy przykładowo wspomnieć odkrycie zjawiska indukcji elektromagnetycznej, stanowiącego podstawę dla rozwoju elektrotechniki, badania teoretyczne Maxwella i doświadczalne Hertza, dotyczące fal elektromagnetycznych, związanych tak ściśle z telekomunikacją itp.

W miarę rozwoju wiedzy ludzkiej fizyczne metody badań znajdują coraz szersze zastosowanie w innych naukach przyrodniczych. Powstają takie dyscypliny naukowe, jak chemia fizyczna, astrofizyka, geofizyka, biofizyka, fizyka ciała stałego itp.

Fizyka jest nauką ścisłą: prawa fizyczne rządzące zjawiskami zachodzącymi w otaczającym nas świecie podawane są w postaci matematycznych zależności. W fizyce doświadczalnej do poznania tych praw prowadzą obserwacja i doświadczenie, w fizyce teoretycznej — odpowiednio interpretowane wyniki analizy matematycznej.

Stosowanie w fizyce naukowych metod badawczych o charakterze doświadczalnym datuje się od czasów Galileusza (wiek XVI). W przeciwieństwie do metod obserwacyjnych, w których badacz tylko rejestruje wyniki obserwacji, w metodach doświadczalnych rola badacza jest czynna. Stwarza on mianowicie określone warunki decydujące o przebiegu zjawiska, wywołuje jego powstanie, wykonuje pomiary, powtarza takie badania wielokrotnie, wreszcie — korzystając ze stałości przebiegu zjawiska w danych warunkach — zestawia wyniki pomiarów i z takich zestawień wyciąga wnioski formu-

łując je początkowo w postaci hipotez. Hipotezy o znaczeniu ogólnym, których słuszność zostaje potwierdzona w licznych badaniach, często przyjmują postać praw lub teorii (np. prawo powszechnego ciążenia, prawa Kirchhoffa, teoria falowa i kwantowa światła, teoria względności itp.). Prawa o podstawowym znaczeniu niekiedy nazywane są zasadami (np. zasady dynamiki, zasada zachowania energii itp.)

W wielu przypadkach prawa fizyczne ustalane są w odniesieniu do ciał wyidealizowanych, modelowych. Wprowadzane są takie pojęcia, jak ciało doskonale sztywne, doskonale sprężyste, punkt materialny, gaz doskonały, ciecz doskonała. Prawa fizyczne wyprowadzane w odniesieniu do ciał modelowych mają uproszczoną postać matematyczną. Oczywiście, są one tylko w przybliżeniu spełnione dla ciał rzeczywistych, lecz niezależnie od tego w zastosowaniach technicznych często oddają cenne usługi zapewniając wymagany stopień dokładności.

Rozwój nauki — ulepszanie metod badawczych i odkrywanie nowych zjawisk — prowadzi niejednokrotnie do stwierdzenia, że poprzednio ustalone prawa nie są dokładne, że wymagają zastąpienia ich prawami nowymi lub też wprowadzenia ograniczenia ich stosowalności. Tak np. znane ogólnie równanie wyrażające matematycznie drugą zasadę dynamiki

$$F = ma,$$

gdzie F oznacza siłę, m — masę, a — przyspieszenie, jest równaniem przybliżonym, spełnionym tym dokładniej, im mniejsza jest prędkość poruszającego się ciała. Należy więc pamiętać, że poznane prawa są prawami obowiązującymi na danym etapie rozwoju nauki.

2. Wielkości fizyczne. Układy jednostek. Układ SI

Mianem *wielkości fizycznej* (albo wielkości mierzalnej) metrologia obejmuje każdą mierzalną cechę zjawiska lub ciała. Przy stosowaniu metody doświadczalnej w badaniach fizycznych poszczególne wielkości są mierzone z większą lub mniejszą dokładnością, zawsze jednak z pewnym błędem. Wyniki pomiarów powinny być podawane w ogólnie przyjętych, dokładnie określonych jednostkach. Dokonując przeglądu jednostek służących do wyrażania powszechnie znanych wielkości fizycznych, np. takich, jak długość, masa, ciśnienie, łatwo można się przekonać, że w tej dziedzinie istnieje jeszcze duża dowolność. Tak np. w Europie do wyrażania długości stosuje się około dwudziestu różnych jednostek, a do wyrażania masy — jeszcze więcej. W wyniku braku unifikacji powstawały w przeszłości różne układy jednostek (np. CGS, MKSA, elektrostatyczny CGS, elektromagnetyczny CGS, techniczny i inne). Wzory definiujące poszczególne wielkości miały różną postać, a współczynniki i stałe fizyczne miały różne wartości liczbowe. Utrudniało to oczywiście wymianę informacji oraz szybkie zapoznawanie się z literaturą i dokumentacją.

Wielkim krokiem naprzód w dziedzinie unifikacji jednostek miar było zaproponowanie w roku 1960 przez XI Generalną Konferencję Miar jednolitego międzynarodowego układu jednostek miar, zwanego układem SI (*Système International d'Unités*),

i zalecenie ogólnego jego stosowania. Zalety tego układu wiążą się nie tylko z jego międzynarodowym charakterem. Układ ten jest tak pomyślany, że może być stosowany do wyrażania prawie wszystkich wielkości w różnych dziedzinach wiedzy: ma on zatem ważną cechę uniwersalności.

Aczkolwiek układ SI nie jest jeszcze opracowany we wszystkich szczegółach, wiele państw, a między nimi i Polska*, wydało już akty prawne wprowadzające układ SI jako obowiązujący.

Tabela 1

Wielkości podstawowe i uzupełniające układu SI

Wielkość	Jednostka układu SI	
	Nazwa	Oznaczenie
A. Wielkości podstawowe		
długość	metr	m
masa	kilogram	kg
czas	sekunda	s
natężenie prądu elektrycznego	amper	A
temperatura termodynamiczna	kelwin	K
światłość	kandela	cd
B. Wielkości uzupełniające		
kąt płaski	radian	rad
kąt bryłowy	steradian	sr

Układ SI opiera się na sześciu wielkościach podstawowych i dwóch uzupełniających. Pierwszą grupę tych wielkości stanowią: długość, masa, czas, natężenie prądu elektrycznego, temperatura oraz światłość. Do grupy drugiej należą: kąt płaski i kąt bryłowy. W tabeli 1 podane jest zestawienie wszystkich wymienionych wielkości (które dalej będziemy nazywali bazowymi) wraz z ich jednostkami miary w układzie SI i symbolami. Jak widać, układ SI ma dużo wspólnego z szeroko stosowanymi dawniej układami CGS i MKSA. Układ jednostek mechanicznych CGS opiera się na tych samych co układ SI podstawowych wielkościach mechanicznych, a mianowicie długości, masie i czasie. Inne jednak są w obu układach jednostki podstawowe. Jeszcze więcej wspólnego mają układy MKSA i SI. Układ MKSA pokrywa się całkowicie z układem SI w dziedzinie podstawowych wielkości mechanicznych i elektrycznych oraz ich podstawowych jednostek.

3. Definicje jednostek bazowych układu SI

Znajomość definicji jednostek fizycznych ma w nauce zasadnicze znaczenie, dlatego też niżej zestawiono definicje wszystkich jednostek bazowych, a więc podstawowych i uzupełniających, układu SI.

---

* W Polsce odpowiednie uchwały zostały ogłoszone w Dzienniku Ustaw, rok 1966 i 1970 oraz w Monitorze Polskim, rok 1966 i 1971.

Metr jest długością równą 1 650 763,73 długości fali w próżni ściśle określonego promieniowania monochromatycznego o barwie pomarańczowej, emitowanego przez izotop kryptonu 86*.

Kilogram jest masą międzynarodowego wzorca przechowywanego w Międzynarodowym Biurze Miar w Sèvres pod Paryżem.

Sekunda jest 1/31 556 925,9747 częścią roku zwrotnikowego 1900**.

Amper jest natężeniem nie zmieniającego się prądu elektrycznego, który — płynąc w dwóch równoległych prostoliniowych, nieskończenie długich przewodach o przekroju okrągłym znikomo małym, umieszczonych w próżni w odległości jednego metra jeden od drugiego — wywołuje między tymi przewodami siłę równą $2 \cdot 10^{-7}$ niutona na każdy metr długości przewodu.

Kelwin jest jednostką temperatury termodynamicznej w skali, w której temperatura punktu potrójnego*** wody jest równa dokładnie 273,16K.

Kandela jest światłością, którą ma w kierunku prostopadłym pole równe $\frac{1}{6} \cdot 10^{-5}\,m^2$ powierzchni ciała doskonale czarnego, promieniującego w temperaturze krzepnięcia platyny pod ciśnieniem 101 325 N/m^2****.

Radian jest kątem płaskim zawartym między dwoma promieniami koła, wycinającymi z okręgu tego koła łuk o długości równej promieniowi.

Steradian jest kątem bryłowym o wierzchołku w środku kuli, wycinającym w powierzchni tej kuli pole równe kwadratowi jej promienia.

Do zilustrowania dążeń metrologów do uzyskania możliwie dokładnej definicji rozważanej jednostki wykorzystamy przykład metra. Pomysł wprowadzenia metra jako jednostki długości został wysunięty podczas Wielkiej Rewolucji Francuskiej. Długość wzorcowej sztaby metrowej miała odpowiadać jednej czterdziestomilionowej części południka przechodzącego przez Paryż. Późniejsze pomiary długości południka były dokładniejsze i okazało się, że sporządzona sztaba wzorcowa ma długość nieco różną od zamierzonej. Zamiast budować nowy wzorzec ustalono, że metrem wzorcowym będzie odległość mierzona w temperaturze zera stopni Celsjusza między dwiema określonymi kreskami zaznaczonymi na sztabie wzorcowej, przechowywanej w Międzynarodowym Biurze Miar w Sèvres.

Ponieważ upadła pierwotna koncepcja powiązania długości metra z długością południka paryskiego, więc mogłaby powstać duża trudność w odtworzeniu metra w razie zaginięcia czy zniszczenia sztaby wzorcowej. Postanowiono więc określić jego długość w inny jeszcze sposób, a mianowicie wykorzystując stałość długości fali linii widmowej danego pierwiastka w określonych warunkach. Trudność polegała m. in. na wyborze

* Definicja nieco uproszczona.

** Definicja nieco uproszczona.

*** Punkt potrójny odpowiada stanowi równowagi między fazą stałą, ciekłą i gazową danej substancji (por. § 17.13). W przypadku wody punkt potrójny był dotychczas charakteryzowany ciśnieniem równym 4,6 mm Hg oraz temperaturą 0,01°C. Z porównania z definicją stopnia kelwina wynika, że temperatura 0°C odpowiada wartości 273,15K.

**** Jednostka ciśnienia wymieniona w końcowej części tej definicji wynika z przeliczenia jednej atmosfery fizycznej na jednostkę SI. Temperatura krzepnięcia platyny pod tym ciśnieniem wynosi 2046,15K, czyli 1773°C.

pierwiastka mającego możliwie wąską linię widmową. Brano pod uwagę czerwoną linię kadmu, zieloną linię izotopu rtęci o masie atomowej 198 oraz linie dwóch izotopów kryptonu. Opierając się na wynikach licznych badań przeprowadzonych w różnych krajach ustalono, że międzynarodowym wzorcem długości będzie 1 650 763,73 długości fali linii pomarańczowej kryptonu o masie atomowej 86. Definicja ta pokrywa się z definicją metra w układzie SI.

4. Wielkości pochodne i ich jednostki w układzie SI

W układzie SI, jak zresztą i w każdym innym układzie jednostek, obok wielkości bazowych występują wielkości pochodne. Wielkości pochodne są określane na podstawie równań definicyjnych bezpośrednio lub pośrednio za pomocą wielkości bazowych danego układu. Tak np. prędkość ruchu jednostajnego, zdefiniowana wzorem

$$v = \frac{s}{t},$$

gdzie v oznacza prędkość ruchu jednostajnego, s — drogę, t — czas, jest wielkością pochodną zdefiniowaną bezpośrednio za pomocą wielkości bazowych układu SI, a mianowicie długości (drogi) s i czasu t. Natomiast przyspieszenie ruchu jednostajnie zmiennego, zdefiniowane wzorem

$$a = \frac{\Delta v}{\Delta t},$$

gdzie a oznacza przyspieszenie ruchu jednostajnie zmiennego, Δv — przyrost prędkości, Δt — przyrost czasu, jest wielkością pochodną zdefiniowaną pośrednio za pomocą wielkości bazowych układu SI, gdyż jedna z wielkości występujących w definicji (Δv) jest już wielkością pochodną.

Każdej wielkości pochodnej należącej do układu SI — podobnie jak i każdej wielkości bazowej — przypisuje się jedną i tylko jedną jednostkę SI. Cechą wspólną wszystkich jednostek wielkości pochodnych w układzie SI jest to, że każdą jednostkę pochodną w układzie SI, oznaczoną symbolicznie $[X]$, można przedstawić w postaci

$$[X] = \mathrm{m}^{\alpha} \cdot \mathrm{kg}^{\beta} \cdot \mathrm{s}^{\gamma} \cdot \mathrm{A}^{\delta} \cdot \mathrm{K}^{\varepsilon} \cdot \mathrm{cd}^{\eta} \cdot \mathrm{rad}^{\vartheta} \cdot \mathrm{sr}^{\lambda},$$

przy czym wykładniki potęg mogą przyjmować dowolne wartości.

Jak widać, podana wyżej zależność nie zawiera żadnego współczynnika liczbowego. Wiąże się to z bardzo ważną cechą układu SI, zwaną spójnością. Tę cechę mają zresztą także układy CGS i MKSA. *Spójność układu jednostek* polega na tym, że każdą jednostkę pochodną danego układu otrzymuje się z jednostek bazowych (bezpośrednio lub pośrednio) przez ich mnożenie lub dzielenie bez wprowadzania dodatkowych współczynników liczbowych*.

---

* Bardziej szczegółowe omówienie spójności układów jednostek, jak również stosowanych zapisów równań fizycznych w postaci tzw. wzorów wielkościowych i liczbowych podane jest w Uzupełnieniach, § U.1.

Jednostki pochodne SI można podzielić na takie, które nie mają nazw specjalnych, np. jednostka prędkości — m/s, jednostka przyspieszenia — m/s^2, oraz takie, którym nadano nazwy specjalne, np. jednostka siły — niuton (skrót N, $1\ N \equiv 1\ kg \cdot m/s^2$), jednostka pracy — dżul (skrót J, $1\ J \equiv 1\ kg \cdot m^2/s^2$) itp. Nazwy nadawane jednostkom pochodzą często od nazwisk zasłużonych badaczy, pisane są fonetycznie i oznaczane w skrócie wielkimi początkowymi literami nazwiska (czasem z dodatkiem drugiej małej litery, np. herc — Hz).

5. Jednostki wielokrotne i podwielokrotne układu SI

Tworzenie jednostek wielokrotnych i podwielokrotnych w układzie SI wymaga nieco szerszego omówienia. W odniesieniu do wszystkich jednostek — z wyjątkiem jednostek masy — obowiązuje zasada tworzenia krotności jednostek od jednostki układu SI. Odrębne potraktowanie jednostek masy wiąże się z tym, że w nazwie jednostki masy układu SI, a mianowicie w kilogramie, występuje już przedrostek „kilo". Dla uniknięcia występowania dwóch przedrostków krotności obok siebie wielokrotne i podwielokrotne jednostki masy wyprowadza się z grama.

W większości przypadków jednostki wielokrotne i podwielokrotne tworzy się według zasad systemu dziesiętnego i wtedy stosuje się szereg ustalonych międzynarodowych przedrostków i ich oznaczeń. Zestawienie nazw przedrostków, ich oznaczeń i stosunku do jednostki wyjściowej podaje tab. 2.

Zgodnie z obowiązującym obecnie w Polsce zarządzeniem (Mon. Pol. Nr 74, 1966, poz. 356) wiele jednostek miar nie stanowiących wymienionych w tab. 2 krotności jednostek SI zostało przejściowo dopuszczonych do stosowania jako legalne. Jako przy-

Tabela 2

Przedrostki określające krotności jednostek w układzie dziesiętnym

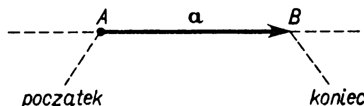
Przedrostek	Oznaczenie	Wielokrotność lub podwielokrotność
tera	T	10^{12}
giga	G	10^{9}
mega	M	10^{6}
kilo	k	10^{3}
hekto	h	10^{2}
deka	da	10^{1}
decy	d	10^{-1}
centy	c	10^{-2}
mili	m	10^{-3}
mikro	μ	10^{-6}
nano	n	10^{-9}
piko	p	10^{-12}
femto	f	10^{-15}
atto	a	10^{-18}

kład może służyć angstrem (1 Å = 10^{-10} m). Na podstawie tego samego zarządzenia dopuszczono przejściowo w odniesieniu do niektórych jednostek stosowanie nazw nie związanych z krotnościami. Tak np. można stosować nazwę tona (1 t = 10^3 kg), aczkolwiek poprawną nazwą jest megagram (Mg), lub nazwę mikron (1 $\mu = 10^{-6}$ m), aczkolwiek poprawną nazwą jest mikrometr (μm). Dla niektórych jednostek dopuszczonych przejściowo do stosowania zastrzeżono zakres stosowalności (np. rok świetlny — w astronomii, mila morska — w żegludze, ar i hektar — do pomiarów powierzchni gruntu itp.). Na podstawie tego samego zarządzenia dozwolone jest stosowanie jednostek czasu: minut, godzin, dni, tygodni, lat itd.

Omówione wyżej zalety układu SI, jak charakter międzynarodowy, uniwersalność i spójność, zdecydowały o uznaniu go za układ prawnie obowiązujący. Jednak w okresie przejściowym zarówno w literaturze naukowej i technicznej, jak i podczas użytkowania dawniej budowanej aparatury pomiarowej często jeszcze można będzie spotkać jednostki tradycyjne. Dlatego też w skład niniejszego podręcznika wchodzą tabele przeliczeniowe ważniejszych jednostek fizycznych dopuszczonych przejściowo do stosowania, a dawniej szeroko stosowanych, na jednostki układu SI (tabele poza tekstem na końcu książki).

6. Podstawowe wiadomości o wektorach

Wielkości fizyczne dzielimy na wielkości kierunkowe (wektorowe), zwane w skróceniu wektorami, i wielkości bezkierunkowe, zwane skalarami. Podczas opisywania wielkości wektorowych powinna być podawana ich bezwzględna wartość liczbowa, zwana też modułem, kierunek, zwrot i punkt przyłożenia. Innymi słowy, *wielkość wektorową* można przedstawić geometrycznie jako odcinek skierowany, tj. odcinek leżący na określonej prostej, mający określony początek i koniec (a więc określony zwrot), jak również określoną długość wyrażającą w pewnej skali bezwzględną wartość danego wektora (rys. 1). W książce niniejszej symbole wielkości wektorowych będą pisane grubą czcionką (np. **a**). Te same symbole wydrukowane czcionką pochyłą cienką oznaczać będą wartości liczbowe (moduły) wektorów (np. *a*). Jako przykłady wektorów mogą służyć: siła, prędkość, przyspieszenie itp.



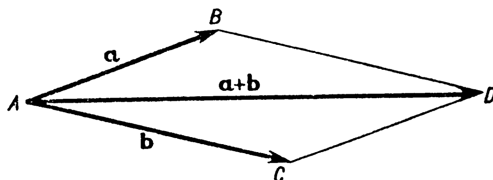
Rys. 1

Skalarami są wielkości, których opis ogranicza się do podania wartości liczbowej. Do skalarów zaliczamy np.: czas, temperaturę, pracę, energię, ładunek elektryczny itp.

Przypomnijmy krótko podstawowe działania wykonywane na wektorach.

Przy dodawaniu wektorów stosuje się tzw. zasadę równoległoboku (rys. 2). Wektor **c** odpowiadający przekątnej równoległoboku, zbudowanego na wektorach **a** i **b**, jest *sumą geometryczną*, czyli *wypadkową* wektorów **a** i **b**.

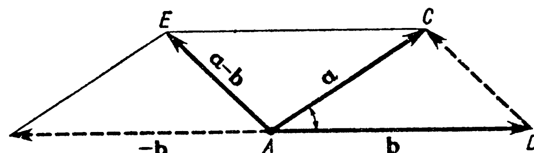
Odejmowanie wektorów **a** i **b** można sprowadzić do dodawania (rys. 3). Do wektora **a** dodajemy wektor —**b**, tzn. wektor co do wartości równy **b**, lecz o zwrocie przeciwnym. Wektor **a**—**b** ma swój początek w punkcie A i odpowiada przekątnej równoległoboku zbudowanego na wektorach **a** i **b**. Z tego samego rysunku widać, że przez bezpośrednie połączenie końców wektorów **a** i **b** wyprowadzonych ze wspólnego punktu A i zaznaczenie zwrotu do odjemnej **a** otrzymujemy również wektor równy co do wartości **a**—**b**, tylko równolegle przesunięty.



Rys. 2

Mnożenie wektora **a** *przez skalar n* daje w wyniku nowy wektor n**a** o wartości liczbowej n razy powiększonej i o zwrocie zgodnym lub przeciwnym względem wektora **a**, zależnie od tego, czy skalar n jest dodatni, czy też ujemny.

Przy *mnożeniu wektora przez wektor* rozróżniamy iloczyn skalarny i iloczyn wektorowy.



Rys. 3

Iloczyn skalarny wektorów **a** i **b** oznaczamy symbolicznie $\mathbf{a} \cdot \mathbf{b}$. *Iloczyn skalarny dwóch wektorów* jest skalarem, którego wartość liczbowa wyraża się iloczynem wartości liczbowych danych wektorów przez cosinus kąta zawartego między nimi, czyli

$$\mathbf{a} \cdot \mathbf{b} = ab\cos\alpha. \tag{1}$$

Jak widać z rys. 4, $OC = b\cos\alpha$ jest rzutem wektora **b** na kierunek wektora **a**. A zatem wartość liczbowa iloczynu skalarnego równa się iloczynowi wartości jednego z wektorów przez rzut na niego wektora drugiego. Z mnożeniem skalarnym wektorów mamy do czynienia w fizyce, np. przy pracy.

Iloczyn wektorowy wektorów **a** i **b** oznaczamy symbolicznie $\mathbf{a} \times \mathbf{b}$. *Iloczyn wektorowy* jest nowym wektorem (rys. 5)

$$\mathbf{a} \times \mathbf{b} = \mathbf{c}$$

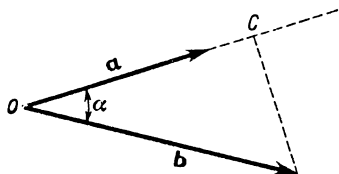
o określonej umownie wartości liczbowej i kierunku.

Wartość liczbowa wektora **c**, czyli c równa się

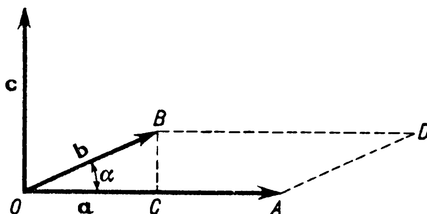
$$c = ab \sin\alpha, \tag{2}$$

gdzie α jest kątem utworzonym przez kierunki wektorów **a** i **b**.

Posługując się rys. 5 łatwo stwierdzić, że iloczyn $ab\sin\alpha$ wyraża pole równoległoboku nakreślonego na wektorach **a** i **b**.

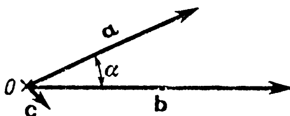


Rys. 4



Rys. 5

Punkt przyłożenia wektora **c** pokrywa się z początkami wektorów **a** i **b**. *Kierunek* jego jest prostopadły do płaszczyzny zawierającej wektory **a** i **b**. *Zwrot wektora* **c** jest określony regułą śruby prawoskrętnej, zwaną również regułą korkociągu. Korkociąg ustawiamy prostopadle do płaszczyzny wektorów **a** i **b** opierając jego ostrze w punkcie O. Rączkę korkociągu ustawiamy równolegle do pierwszego wektora wymienionego w iloczynie wektorowym, a więc w naszym przykładzie do wektora **a**. Obracamy rączkę tak, aby po skręceniu o kąt mniejszy od 180° (w naszym przypadku o kąt α zaznaczony na rys. 5) zajęła ona położenie równoległe do wektora **b**. Podczas tego obrotu ostrze korkociągu przesuwa się w określonym kierunku, który umownie przyjęto za zwrot wektora **c**. Tak np. jeśli na rys. 5 wektory **a** i **b** leżą w płaszczyźnie poziomej, to podczas obrotu ostrze korkociągu przesuwa się w kierunku pionowym w górę. Taki też jest kierunek i zwrot wektora **c**. Gdyby wektory **a** i **b** leżały w płaszczyźnie rysunku (rys. 6), to wektor **c** byłby do niej prostopadły. Ustawieniu wektorów **a** i **b** jak na rys. 6 odpowiada zwrot wektora **c** za płaszczyznę rysunku.

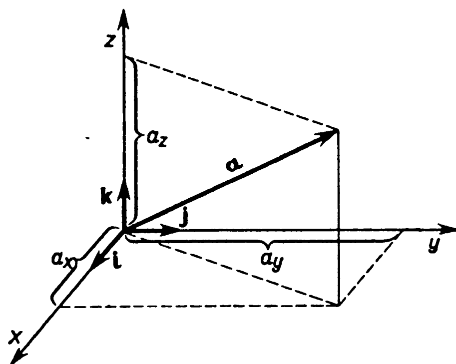


Rys. 6

Zastanówmy się, czy do iloczynu wektorowego stosuje się prawo przemienności, tzn. czy zmiana kolejności wektorów ma wpływ na wartość, kierunek i zwrot wektora **c**. Innymi słowy, chcemy sprawdzić, czy iloczyn wektorowy $\mathbf{a}\times\mathbf{b}$ równa się iloczynowi wektorowemu $\mathbf{b}\times\mathbf{a}$. Wartość liczbowa $ab\sin\alpha$ nie zależy od kolejności czynników i zawsze przedstawia pole równoległoboku zbudowanego na wektorach **a** i **b**. Kierunek iloczynu wektorowego pozostaje prostopadły do płaszczyzny wektorów **a** i **b**, ale zwrot ulega zmianie na przeciwny. Obracając rączkę korkociągu (rys. 5) od położenia równoległego do wektora **b** do położenia równoległego do wektora **a** (o kąt mniejszy od 180°)

wywołamy przesunięcie korkociągu pionowo w dół. Stąd wniosek, że *zmiana kolejności wektorów w iloczynie wektorowym nie zmienia jego wartości liczbowej, lecz zmienia zwrot*. Innymi słowy, $\mathbf{a}\times\mathbf{b} = -\mathbf{b}\times\mathbf{a}$.

Pojęcie iloczynu wektorowego spotkamy wielokrotnie w kursie fizyki. Wymienimy dla przykładu moment siły, moment pędu, siłę Coriolisa, siłę elektrodynamiczną itp.



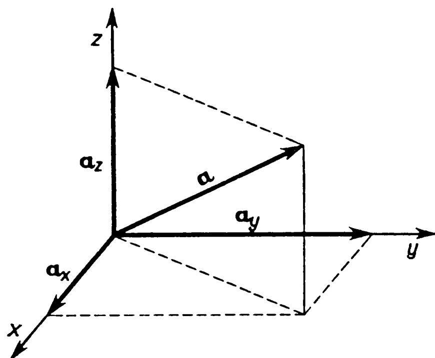
Rys. 7

Na zakończenie wspomnimy jeszcze o sposobie przedstawiania badanego wektora **a** za pomocą jego rzutów na osie współrzędnych prostokątnych (x, y, z). Niech długości tych rzutów będą odpowiednio a_x, a_y i a_z (rys. 7). Wprowadźmy wektory jednostkowe **i, j, k** odpowiednio w kierunku osi x, y, z. Wtedy wektor a może być przedstawiony jako suma wektorów składowych

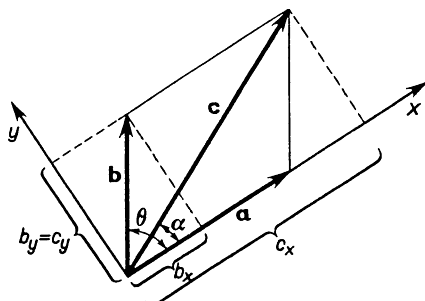
$$\mathbf{a} = a_x\mathbf{i}+a_y\mathbf{j}+a_z\mathbf{k}. \tag{3}$$

Oznaczając składowe wektora a o kierunkach trzech osi współrzędnych (rys. 8) odpowiednio przez

$$a_x\mathbf{i} = \mathbf{a}_x, \quad a_y\mathbf{j} = \mathbf{a}_y, \quad a_z\mathbf{k} = \mathbf{a}_z$$



Rys. 8



Rys. 9

otrzymamy zapis skrócony

$$\mathbf{a} = \mathbf{a}_x + \mathbf{a}_y + \mathbf{a}_z. \tag{4}$$

Ten analityczny sposób przedstawiania wektorów jest bardzo wygodny, gdyż upraszcza wiele działań na wektorach.

P r z y k ł a d 1. Wykazać, że w wyniku sumowania wektorów **a** i **b** tworzących kąt θ otrzymuje się nowy wektor **c** taki, że jego rzuty na prostokątne osie x i y spełniają zależności

$$c_x = a_x + b_x, \quad c_y = a_y + b_y. \tag{5}$$

Wybierzmy układ współrzędnych xy (rys. 9) tak, by jego początek zgadzał się z początkiem obu wektorów i jedna z osi, np. oś x, była zgodna z wektorem **a**. Stosując zasadę równoległoboku znajdujemy wektor **c**. Jego rzuty na osie współrzędnych są odpowiednio równe c_x i c_y. Proste rozważania geometryczne wykazują, że rzeczywiście c_x i c_y spełniają warunek (5). Z rysunku też widać, że nachylenie wektora wypadkowego **c** względem osi x (a zatem także względem wektora **a**) można wyrazić za pomocą zależności

$$\operatorname{tg}\alpha = \frac{c_y}{c_x}.$$

Podany dowód dotyczy szczególnego wyboru położenia osi współrzędnych. Czytelnikowi pozostawiamy rozpatrzenie przypadku ogólnego.

P r z y k ł a d 2. W odniesieniu do wektorów **a** i **b** wymienionych w przykładzie 1 wykazać, opierając się na rozkładzie na składowe, że wartość liczbowa $c = \sqrt{a^2+b^2+2ab\cos\theta}$.

Rzuty danych wektorów na osie wynoszą odpowiednio

$$a_x = a, \quad b_x = b\cos\theta,$$

$$a_y = 0, \quad b_y = b\sin\theta.$$

A zatem zgodnie z (5)

$$c_x = a + b\cos\theta, \quad c_y = b\sin\theta.$$

Stosując twierdzenie Pitagorasa otrzymamy

$$c^2 = c_x^2 + c_y^2 = (a+b\cos\theta)^2 + b^2\sin^2\theta,$$

skąd

$$c = \sqrt{a^2+b^2+2ab\cos\theta}.$$

W dalszych naszych rozważaniach przekonamy się, że z punktu widzenia swobody wyboru punktu przyłożenia można wektory podzielić na trzy grupy, a mianowicie na: 1) wektory związane, 2) wektory przesuwalne, czyli osiowe i 3) wektory swobodne. Punkt przyłożenia wektorów związanych jest dokładnie określony. Jako przykład mogą służyć prędkości lub przyspieszenie punktu materialnego, siła działająca na punkt materialny itd. Punkt przyłożenia wektorów przesuwalnych można wybierać dowolnie na prostej ich działania. Jako przykład można wymienić siłę działającą na ciało sztywne. Punkt przyłożenia wektorów swobodnych może być dowolnie wybierany. Przykładem takiego wektora jest moment pary sił.

Zakończenie niniejszego rozdziału, jak zresztą i następnych, stanowi zestawienie szeregu pytań i zadań nawiązujących do najważniejszych punktów treści i pomyślanych jako pomoc przy przyswajaniu materiału. Bardzo wskazane jest staranne opracowanie odpowiedzi z dużą dbałością o ścisłość merytoryczną.

Pytania i zadania

1. Co jest przedmiotem fizyki?
2. Jakimi metodami posługuje się fizyka doświadczalna?
3. Jaka jest rola ciał modelowych w fizyce?
4. Dlaczego wprowadzenie układu SI jest wielkim krokiem naprzód w dziedzinie metrologii?
5. Na jakich wielkościach podstawowych i uzupełniających opiera się układ SI?
6. Podać zasady tworzenia jednostek wielokrotnych i podwielokrotnych w układzie SI.
7. Na czym polega spójność układu SI?
8. Podać definicje wielkości skalarnych i wektorowych.
9. Omówić podstawowe działania na wektorach: a) dodawanie, b) odejmowanie, c) mnożenie i dzielenie przez skalary dodatnie i ujemne, d) mnożenie skalarne dwóch wektorów, e) mnożenie wektorowe dwóch wektorów.
10. Korzystając z rozkładu wektorów na składowe skierowane wzdłuż osi układu prostokątnego wykazać, że iloczyn skalarny

$$\mathbf{a}\cdot\mathbf{b} = a_x b_x + a_y b_y.$$

11. Jakie są wartości liczbowe iloczynów $\mathbf{a}\cdot\mathbf{a}$ i $\mathbf{a}\times\mathbf{a}$?
12. W układzie prawoskrętnym xyz znajdują się dwa wektory: wektor **a** zgodny z kierunkiem osi x, równy 20 jednostkom oraz wektor **b**, leżący w płaszczyźnie xy, tworzący kąt 30° z osią x, równy 10 jednostkom. Znaleźć iloczyny skalarne $\mathbf{a}\cdot\mathbf{b}$ i $\mathbf{b}\cdot\mathbf{a}$ oraz iloczyny wektorowe $\mathbf{a}\times\mathbf{b}$ i $\mathbf{b}\times\mathbf{a}$.
13. Punkt materialny wykonuje cztery kolejne przesunięcia w jednej płaszczyźnie: 1) 2 m na północ, 2) 4 m na północny wschód, 3) 1 m na wschód, 4) 4 m na południowy zachód. Niech osie współrzędnych wskazują: x — na wschód, y — na północ. Obliczyć: a) składowe wszystkich wektorów na osiach współrzędnych, b) składowe przesunięcia wypadkowego, c) wartość liczbową i kierunek przesunięcia wypadkowego.

CZĘŚĆ I

Mechanika

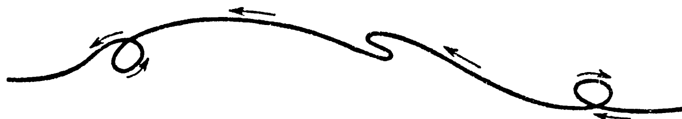
ROZDZIAŁ 1

Kinematyka punktu materialnego i bryły sztywnej

1.1. Określenie ruchu

Kinematyka podaje opis ruchu bez uwzględniania przyczyn i warunków, w jakich dany ruch powstaje. Przez *ruch ciała* rozumiemy zmianę jego położenia w stosunku do innych ciał, które uważamy za nieruchome. Ciała te nazywamy *układem odniesienia.* Tak np. ruch pociągu można rozpatrywać względem ciał w stosunku do Ziemi nieruchomych: budynku stacyjnego, drzew itp. W tym przypadku nie bierzemy pod uwagę ruchu budynku lub drzew, odbywanego wraz z Ziemią dokoła jej osi i dokoła Słońca.

Układy odniesienia można dobierać dowolnie, ale opis ruchu może zależeć od wyboru tego układu. Dobitny przykład tej zależności występuje przy rozpatrywaniu ruchu planet, gdy kolejno Ziemię i Słońce przyjmujemy za układ odniesienia. W pierwszym przypadku ruch jest skomplikowany, odbywa się po liniach krzywych tworzących pętle (rys. 1.1). Ruch tej samej planety odniesiony do Słońca jest znacznie prostszy. Zgodnie z prawem Keplera (por. § 4.3) odbywa się po elipsie. W jednym z ognisk tej elipsy znajduje się Słońce.



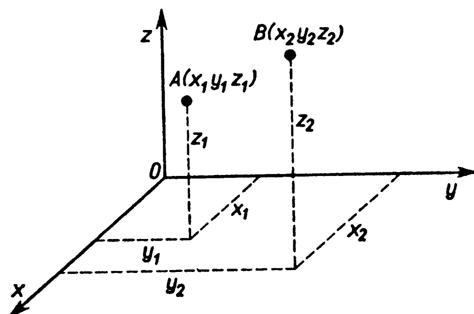
Rys. 1.1

Wstępne rozważania kinematyczne dotyczą ruchu tzw. punktu materialnego. Przez *punkt materialny* rozumiemy ciało modelowe, fikcyjne, obdarzone pewną masą, o rozmiarach takich, że podczas rozważania danego ruchu można je zaniedbać.

Podczas swego ruchu punkt materialny przesuwa się do coraz innych punktów przestrzeni. Zbiór tych punktów stanowi *tor ruchu,* który może być linią prostą lub krzywą. W zależności od kształtu toru mamy do czynienia z *ruchem prostoliniowym*

lub *krzywoliniowym*. Ruch krzywoliniowy może być płaski lub przestrzenny. Przykładem ruchu płaskiego jest ruch po elipsie, po paraboli, po okręgu itp., przykładem ruchu przestrzennego (trójwymiarowego) — ruch po linii śrubowej. Długość przebytego odcinka toru stanowi *drogę ciała*.

Zastanówmy się nad sposobem matematycznego ujęcia ruchów ciał. Wiemy, że położenie punktu w przestrzeni można określać przez podanie współrzędnych w umówionym układzie, np. w układzie współrzędnych kartezjańskich (rys. 1.2). Położenie punktu A w przestrzeni jest określone, gdy znamy jego współrzędne (x_1, y_1, z_1). Podobnie położenie punktu B jest określone, gdy znamy jego współrzędne (x_2, y_2, z_2) w tym samym układzie.



Rys. 1.2

Wyobraźmy sobie, że punkt materialny w swoim ruchu przesuwa się po pewnym torze od położenia A do położenia B. Będziemy mogli twierdzić, że znamy ruch tego punktu, jeśli znane nam będą zmiany współrzędnych x, y, z w zależności od czasu, czyli innymi słowy, jeśli będziemy znali x, y, z jako funkcje czasu t:

$$x = f_1(t), \quad y = f_2(t), \quad z = f_3(t).$$

Funkcje te w zależności od charakteru będą przedstawiały różne rodzaje ruchu. Przed przystąpieniem do szczegółowego omówienia kilku z nich trzeba zdefiniować zasadnicze wielkości kinematyczne.

1.2. Prędkość i przyspieszenie

Podstawowymi wielkościami kinematycznymi są prędkość i przyspieszenie. Korzystając ze znajomości rachunku różniczkowego można je zdefiniować jako odpowiednie pochodne drogi względem czasu. *Prędkość* v jest pierwszą pochodną drogi s względem czasu t:

$$v = \frac{ds}{dt}. \tag{1.1}$$

Przyspieszenie a jest pierwszą pochodną prędkości względem czasu lub drugą pochodną drogi względem czasu:

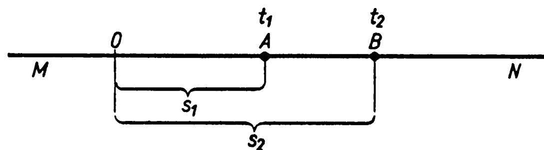
$$a = \frac{dv}{dt} = \frac{d^2s}{dt^2}. \tag{1.2}$$

Prędkość i przyspieszenie zdefiniowane wzorami (1.1) i (1.2) określają ruch punktu materialnego w danej chwili, a więc są wielkościami chwilowymi.

Z podanych definicji wynikają jednostki obu wielkości. W układzie SI są nimi odpowiednio m/s oraz m/s^2.

Niejednokrotnie na początku studiów występują trudności w kojarzeniu sensu fizycznego wzoru z jego matematycznym ujęciem. W związku z tym rozszerzymy niniejszy punkt przypominając elementarne ujęcie pojęć prędkości średniej i chwilowej oraz przyspieszenia średniego i chwilowego.

Prędkość średnia. W chwili $t = t_1$ ciało znajduje się w punkcie A (rys. 1.3), czyli jego droga względem punktu O wynosi s_1. W chwili $t = t_2$ ciało znajduje się w punkcie B, czyli jego droga względem punktu O wynosi s_2. A zatem w czasie t_2-t_1 ciało przebyło drogę s_2-s_1.



Rys. 1.3

Stosunek długości przebytej drogi s_2-s_1 do czasu t_2-t_1, w którym została ona przebyta, nosi nazwę ***prędkości średniej w przedziale czasu od t_1 do t_2***.

Prędkość średnią oznaczamy symbolem $\bar{v}$. A więc

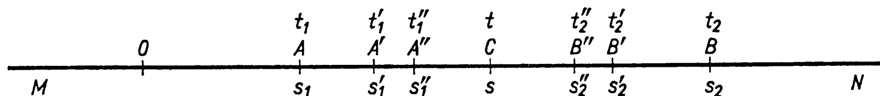
$$\bar{v} = \frac{s_2-s_1}{t_2-t_1} = \frac{\Delta s}{\Delta t}. \tag{1.3}$$

Symbol Δ jest użyty do oznaczenia skończonych przyrostów danych wielkości.

Prędkość średnia charakteryzuje ruch zmienny w sposób przybliżony. Wyraża ona prędkość, jaką posiadałoby dane ciało, gdyby przebywało drogę Δs w czasie Δt ruchem jednostajnym.

Prędkość chwilowa. Prędkość średnia na odcinku s_2-s_1 nie określa prędkości, z jaką ciało badane mija dowolny punkt tego odcinka. Znając prędkość średnią na odcinku s_2-s_1, czyli w przedziale czasu od t_1 do t_2, nic jeszcze nie można powiedzieć o wartości prędkości w pewnej określonej chwili t zawartej między t_1 i t_2.

Do określenia prędkości chwilowej, zwanej również rzeczywistą lub po prostu prędkością w danym punkcie toru, doprowadza rozumowanie następujące. Niech torem punktu będzie prosta MN (rys. 1.4). W chwili t_1 ciało mija punkt A odległy od punktu



Rys. 1.4

O o s_1, w chwili t_2 — punkt B odległy od punktu O o s_2, a zatem prędkość średnia na odcinku s_2-s_1 wynosi

$$\bar{v} = \frac{s_2-s_1}{t_2-t_1}.$$

Chcemy określić prędkość ciała w chwili t przy mijaniu punktu C. W tym celu nie poprzestajemy na ustaleniu dróg s_1 i s_2, odpowiadających czasom t_1 i t_2, lecz zwiększamy liczbę pomiarów zbliżając stopniowo punkty A i B do punktu C. Stwierdzamy, że stopniowo wzrastającym czasom $t_1', t_1'', \ldots$ odpowiadają stopniowo wzrastające drogi $s_1', s_1'', \ldots$, jak również stopniowo malejącym czasom $t_2', t_2'', \ldots$ odpowiadają stopniowo malejące drogi $s_2', s_2'', \ldots$ Z tych wielkości możemy utworzyć wyrażenia typu

$$\frac{s_2'-s_1'}{t_2'-t_1'}, \quad \frac{s_2''-s_1''}{t_2''-t_1''}, \ldots$$

Każde z tych wyrażeń przedstawia prędkość średnią na odcinku coraz krótszym, zawierającym punkt C. Moglibyśmy utworzyć nieskończony ciąg wyrażeń tego typu przy spełnieniu warunków

$$t_1' < t_1'' < \ldots < t$$

oraz

$$t_2' > t_2'' > \ldots > t,$$

z których wynika, że oba czasy t_1 i t_2 zmierzają do wspólnej wartości t, odpowiadającej mijaniu punktu C. Oczywiście, równocześnie

$$s_1' < s_1'' < \ldots < s$$

oraz

$$s_2' > s_2'' > \ldots > s,$$

czyli drogi zmierzają do wspólnej wartości s, odpowiadającej położeniu punktu C. Stwierdzilibyśmy wtedy, że ten ciąg nieskończony ma określoną granicę. Warunek, że czasy $t_1^{(n)}$ i $t_2^{(n)}$ zmierzają do wspólnej granicy t jest równoznaczny z warunkiem, że punkty A i B zbliżają się do punktu C, a zatem, że różnica czasów $\Delta t = t_2-t_1$ zmierza do zera.

Granicę ciągu prędkości średnich, gdy Δt zmierza do zera, nazywamy *prędkością chwilową* albo po prostu *prędkością w chwili t lub w punkcie C toru*:

$$v = \lim_{\Delta t\to 0} \bar{v} = \lim_{\Delta t\to 0} \frac{s_2-s_1}{t_2-t_1} = \lim_{\Delta t\to 0} \frac{\Delta s}{\Delta t}. \tag{1.4}$$

Znając elementy rachunku różniczkowego można to wyrażenie zastąpić symbolem pochodnej drogi względem czasu

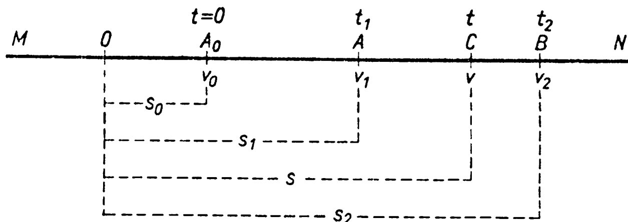
$$v = \frac{ds}{dt}.$$

Przyspieszenie średnie i chwilowe. Niech w chwili t_1 przy mijaniu punktu A (rys. 1.5) ciało ma prędkość v_1, a w chwili t_2 przy mijaniu punktu B — prędkość v_2. W czasie t_2-t_1 prędkość uległa zmianie o v_2-v_1. Na jednostkę czasu przypada zatem średni przyrost prędkości równy $(v_2-v_1)/(t_2-t_1)$. Ten stosunek nazywamy *średnim przyspie-*

szeniem na odcinku drogi między punktami A i *B w przedziale czasu od t_1 do t_2*. Przyspieszenie średnie oznaczamy symbolem $\bar{a}$:

$$\bar{a} = \frac{v_2-v_1}{t_2-t_1} = \frac{\Delta v}{\Delta t}. \tag{1.5}$$

Zwężając granice przedziału czasu t_2-t_1 można — w sposób analogiczny do omówionego w poprzednim punkcie — utworzyć nieskończony ciąg wartości przyspieszeń



Rys. 1.5

średnich. Gdy czasy t_1 i t_2 zmierzają do wspólnej granicznej wartości t, ciąg wartości przyspieszeń średnich zmierza do granicy, którą nazywamy *przyspieszeniem chwilowym w chwili t* (*w punkcie C*) lub — krótko — przyspieszeniem i oznaczamy literą a:

$$a = \lim_{\Delta t \to 0} \frac{v_2-v_1}{t_2-t_1} = \lim_{\Delta t \to 0} \frac{\Delta v}{\Delta t}. \tag{1.6}$$

Stosując symbolikę rachunku różniczkowego można napisać

$$a = \frac{dv}{dt} = \frac{d^2s}{dt^2}.$$

1.3. Ruch prostoliniowy jednostajny i jednostajnie zmienny

W § 1.1 ustaliliśmy, że w przypadku ogólnym opis ruchu punktu materialnego w przestrzeni wymaga znajomości współrzędnych x, y, z jako funkcji czasu t. Matematyczny opis ruchu jest najprostszy, gdy torem punktu materialnego jest linia prosta. Niech prosta MN (rys. 1.3) przedstawia tor punktu materialnego. Punkt O, dowolnie wybrany na tej prostej, niech będzie punktem, względem którego liczymy przebyte drogi. Jeśli w chwili t_1 punkt materialny znajduje się w punkcie A, to jego drogą jest odległość $OA = s_1$. Jeśli w chwili t_2 punkt materialny znajduje się w punkcie B, to jego drogą jest $OB = s_2$ itd. Ruch prostoliniowy po danym torze jest znany, gdy wiemy, jaką funkcją czasu jest droga s, czyli gdy znamy

$$s = f(t).$$

Zależności drogi od czasu mogą być różne. Przykładowo wymienimy

$$s = Bt+C, \tag{1.7}$$

$$s = Bt^2+Ct+D, \tag{1.8}$$

$$s = A\sin(\omega t). \tag{1.9}$$

Współczynniki A, B, C, D i ω są wielkościami stałymi, tzn. niezależnymi od czasu.

Analiza wyżej podanych zależności pozwoli określić cechy fizyczne odpowiednich ruchów. Zbadajmy prędkość i przyspieszenie w ruchach (1.7) i (1.8), odkładając analizę ruchu (1.9) do rozdz. 10.

Różniczkując względem czasu wyrażenie (1.7) znajdujemy

$$v = \frac{ds}{dt} = B = \text{const},$$

czyli prędkość w ruchu badanym jest niezależna od czasu. Jest to zatem dobrze znany *ruch jednostajny*. Przyspieszenie w tym ruchu równa się

$$a = \frac{dv}{dt} = 0.$$

Zastanówmy się jeszcze nad znaczeniem fizycznym drugiej stałej występującej w równaniu (1.7), a mianowicie stałej C. Oznaczmy przez s_0 drogę odpowiadającą początkowi rachuby czasu, czyli chwili $t = 0$ (rys. 1.5). Z ogólnego wzoru na drogę wynika, że

$$s_0 = C.$$

Stała C oznacza więc drogę początkową s_0, odpowiadającą początkowi rachuby czasu. A zatem ruch o stałej prędkości v, czyli ruch jednostajny, opisuje równanie

$$s = vt + s_0. \tag{1.10}$$

Analogicznie badamy ruch prostoliniowy określony równaniem

$$s = Bt^2 + Ct + D.$$

Prędkość w tym ruchu wynosi

$$v = \frac{ds}{dt} = 2Bt + C. \tag{1.11}$$

Niech w chwili $t = 0$, a więc w chwili, od której rozpoczynamy rachubę czasu, ciało ma prędkość początkową v_0. Wtedy

$$v_0 = C.$$

W ten sposób określiliśmy znaczenie stałej C.

Przyspieszenie w tym ruchu znajdujemy korzystając z (1.11):

$$a = \frac{dv}{dt} = 2B = \text{const}.$$

Jak widać, przyspieszenie w ruchu badanym jest stałe, mamy więc do czynienia z ruchem jednostajnie zmiennym.

Ustalmy wartości pozostałych współczynników stałych B i D. Z ostatniego wzoru wynika, że

$$B = \frac{a}{2}.$$

Znaczenie stałej D znajdujemy wprowadzając drogę początkową s_0, odpowiadającą czasowi $t = 0$:

$$s_0 = D.$$

Ostatecznie więc wzory na drogę i prędkość w ruchu prostoliniowym jednostajnie zmiennym przyjmują postać

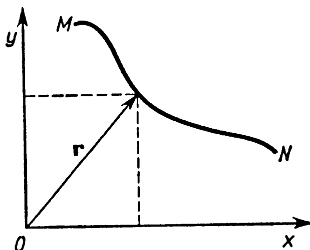
$$s = s_0 + v_0 t + \frac{at^2}{2}, \tag{1.12}$$

$$v = v_0 + at. \tag{1.13}$$

Rozważając takie wielkości jak *prędkość i przyspieszenie* należy pamiętać, że *są to wielkości wektorowe*. W ruchu prostoliniowym prędkość jest oczywiście skierowana wzdłuż toru. Prędkość w ruchu prostoliniowym jednostajnym jest stała zarówno co do wartości, jak i co do kierunku. W ruchu prostoliniowym zmiennym kierunek prędkości jest również stały, ale jej wartość liczbowa ulega zmianie. W ruchu tym występuje przyspieszenie mające zwrot zgodny ze zwrotem prędkości w ruchu przyspieszonym, a przeciwny w ruchu opóźnionym (wtedy czasem jest nazywane opóźnieniem). Tak np. w rzucie pionowym do góry prędkość jest stale skierowana pionowo w górę, a przyspieszenie (ziemskie) — stale pionowo w dół. Podczas swobodnego spadku ciał prędkość i przyspieszenie skierowane są w dół.

1.4. Ruch krzywoliniowy

Stopniując trudności w rozważaniach kinematycznych przejdziemy od ruchu prostoliniowego do ruchu krzywoliniowego płaskiego, a następnie przestrzennego. Do opisu toru ruchu krzywoliniowego płaskiego w układzie współrzędnych (x, y) lub przestrzennego w układzie (x, y, z) może służyć *wektor promienia wodzącego* $\mathbf{r}$, którego początek leży stale w początku układu współrzędnych, a koniec stale zmienia swoje poło-



Rys. 1.6

żenie przesuwając się wzdłuż toru ruchu (rys. 1.6). Ruch krzywoliniowy jest opisany, gdy znamy $\mathbf{r} = f(t)$, tzn. gdy wiemy, jak z biegiem czasu zmienia się wartość liczbowa r oraz jego kierunek. Ale jak wiadomo, każdy wektor można przedstawić za pomocą jego rzutów na wybrane osie. W ruchu krzywoliniowym płaskim

$$\mathbf{r} = x\mathbf{i}+y\mathbf{j},$$

gdzie **i** i **j** są znanymi nam już jednostkowymi wektorami w kierunku osi x i y.

Analogicznie, w ruchu krzywoliniowym przestrzennym

$$\mathbf{r} = x\mathbf{i}+y\mathbf{j}+z\mathbf{k}.$$

Oczywiście, ruch płaski można traktować jako przypadek szczególny ruchu przestrzennego.

Jeżeli wiemy, jak z biegiem czasu zmieniają się x, y i z, czyli znamy funkcje $x = x(t)$, $y = y(t)$ i $z = z(t)$, to wiemy też, jak z biegiem czasu zmienia się **r**, czyli znamy ruch.

W praktyce najczęściej badanie toru ruchu krzywoliniowego sprowadza się do wyznaczenia funkcji $f(x, y, z) = 0$, którą otrzymuje się z funkcji $x(t)$, $y(t)$ i $z(t)$ przez wyrugowanie czasu t. Dalsze badanie ruchu krzywoliniowego obejmuje wyznaczanie prędkości i przyspieszenia. Omówimy to bardziej szczegółowo w następnych punktach, zanim jednak do tego przejdziemy, przypomnimy tzw. zasadę niezależności ruchów.

Według zasady niezależności ruchów, jeśli punkt materialny bierze udział równocześnie w kilku ruchach (ruch złożony), to każdy z tych ruchów (składowych) odbywa się bez zakłóceń tak, jakby pozostałych ruchów nie było.

A zatem w ruchu złożonym obowiązują zależności następujące:

$$\mathbf{v} = \mathbf{v}_1+\mathbf{v}_2+ \ldots,$$

$$\mathbf{a} = \mathbf{a}_1+\mathbf{a}_2+ \ldots,$$

tzn. *prędkość* **v** *ruchu złożonego* jest wypadkową prędkości $\mathbf{v}_1$, $\mathbf{v}_2$ ... ruchów składowych i *przyspieszenie* **a** *ruchu złożonego* jest wypadkową przyspieszeń $\mathbf{a}_1$, $\mathbf{a}_2$ ... ruchów składowych.

Ogólnie znane są liczne przykłady ruchów złożonych, np. ruch względem Ziemi pasażera idącego po pokładzie płynącego statku, ruch względem Ziemi łodzi popychanej siłą wioseł i unoszonej przez prąd wody, ruch satelity ziemskiego wynikający z nadanej mu prędkości początkowej i z istnienia przyspieszenia grawitacyjnego itp.

W dalszych rozważaniach będziemy wykorzystywali zasadę niezależności ruchów zarówno wtedy, gdy ruch złożony będziemy zastępowali ruchami składowymi w kierunku osi współrzędnych (np. w punktach 1.4.2 i 1.4.3), jak i wtedy, gdy mając dane ruchy składowe w kierunku osi współrzędnych, będziemy szukali ruchu wypadkowego (np. w punktach 1.4.1, 1.4.4 i w § 1.5). Szczegółowo omówimy dwa przykłady ruchów złożonych, a mianowicie rzut poziomy i rzut ukośny, w § 4.7.

1.4.1. *Wyznaczanie równania toru w ruchu płaskim.* Zasadę wyznaczania toru poznaliśmy już w poprzednim paragrafie. Obecnie zastosujemy tę zasadę do następującego przypadku.

P r z y k ł a d. Określić równanie toru ruchu płaskiego, w którym

$$x = A\cos\omega t, \tag{1.14}$$

$$y = B\sin\omega t, \tag{1.15}$$

gdzie A i B są to wielkości stałe, niezależne od czasu.

Równanie toru znajdujemy rugując z powyższych zależności czas t. Z (1.14) wynika, że

$$\cos\omega t = \frac{x}{A},$$

skąd

$$\cos^2\omega t = \frac{x^2}{A^2}, \quad \text{czyli} \quad 1-\sin^2\omega t = \frac{x^2}{A^2}.$$

Uwzględniając (1.15) otrzymujemy

$$1-\frac{y^2}{B^2} = \frac{x^2}{A^2},$$

czyli

$$\frac{x^2}{A^2}+\frac{y^2}{B^2} = 1. \tag{1.16}$$

Jest to ogólnie znane równanie elipsy o półosiach A i B (rys. 1.7). W przypadku gdy $A = B$, otrzymuje się równanie okręgu:

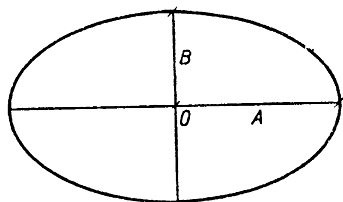
$$x^2+y^2 = A^2. \tag{1.17}$$

A zatem w rozważanym przykładzie torem ruchu jest elipsa przechodząca w okręg przy narzuceniu dodatkowego warunku, że $A = B$.

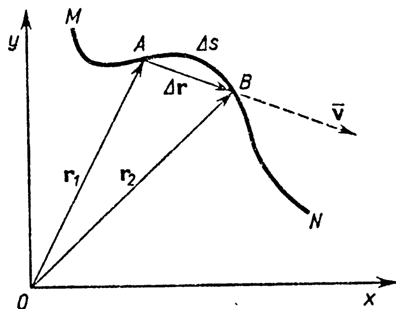
P y t a n i e. Po jakim torze posuwa się punkt materialny, jeśli

$$x = Bt, \quad y = Ct,$$

przy czym B i C są wielkościami stałymi, niezależnymi od czasu?



Rys. 1.7



Rys. 1.8

1.4.2. *Badanie prędkości w ruchu krzywoliniowym.* Rozpatrzmy ruch krzywoliniowy płaski przedstawiony na rys. 1.8. Niech wektory $\mathbf{r}_1$ i $\mathbf{r}_2$ przedstawiają promienie wodzące odpowiednio w chwilach t_1 i t_2. Wektor $\Delta\mathbf{r} = \mathbf{r}_2-\mathbf{r}_1$ jest przyrostem wektora promienia wodzącego w czasie $\Delta t = t_2-t_1$. Długość tego wektora tym mniej różni się od długości drogi Δs przebytej wzdłuż toru, im krótszy jest czas obserwacji Δt. Wyrażenie

$$\bar{\mathbf{v}} = \frac{\Delta\mathbf{r}}{\Delta t}$$

przedstawia *wektor prędkości średniej*, zgodny co do kierunku z $\Delta\mathbf{r}$. Zakładając, że $\Delta t \to 0$ otrzymujemy *prędkość wektorową chwilową* w punkcie A toru

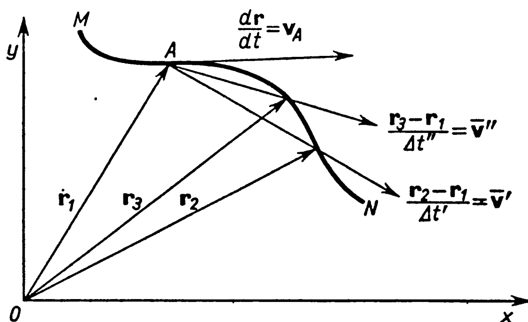
$$\mathbf{v} = \lim_{\Delta t\to 0}\frac{\Delta\mathbf{r}}{\Delta t} = \frac{d\mathbf{r}}{dt}.$$

Wartość liczbowa wektora **v** wynika z następujących równości:

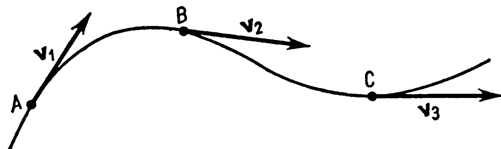
$$v = \lim_{\Delta t \to 0} \frac{\Delta r}{\Delta t} = \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t} = \frac{ds}{dt},$$

gdyż $\Delta r \to \Delta s$, gdy $\Delta t \to 0$.

Kierunek wektora prędkości chwilowej jest styczny do toru w danym punkcie. Widać to wyraźnie z rys. 1.9. Zmniejszając stopniowo czasy $\Delta t'$, $\Delta t''$, ... itd. otrzymamy ciąg wektorów $\bar{\mathbf{v}}'$, $\bar{\mathbf{v}}''$, ... itd. stanowiących prędkości średnie, skierowanych wzdłuż siecznych. Granicznym położeniem tych siecznych, gdy $\Delta t \to 0$, jest styczna do krzywej poprowadzona w punkcie A. Stąd wniosek, że kierunek prędkości chwilowej w danym punkcie toru krzywoliniowego jest styczny do krzywej w danym punkcie (rys. 1.10).



Rys. 1.9



Rys. 1.10

Wektor prędkości **v**, podobnie jak wektor **r**, można przedstawić za pomocą składowych skierowanych wzdłuż osi współrzędnych. A mianowicie, jeżeli — jak poprzednio —

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j},$$

to

$$\mathbf{v} = \frac{d\mathbf{r}}{dt} = \frac{dx}{dt}\mathbf{i} + \frac{dy}{dt}\mathbf{j},$$

czyli

$$\mathbf{v} = v_x\mathbf{i} + v_y\mathbf{j},$$

gdzie

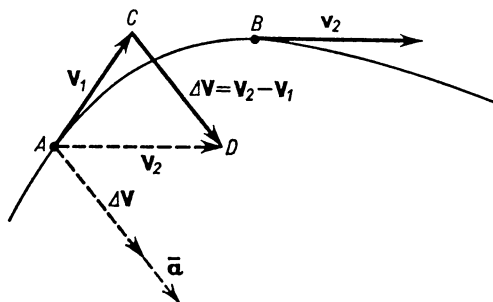
$$v_x = \frac{dx}{dt}, \quad v_y = \frac{dy}{dt}.$$

W przypadku ruchu przestrzennego otrzymamy wyrażenie ogólniejsze

$$\mathbf{v} = v_x\mathbf{i} + v_y\mathbf{j} + v_z\mathbf{k},$$

gdzie $v_z = dz/dt$. A zatem wektor prędkości w danym punkcie toru ruchu krzywoliniowego ma kierunek styczny do toru w badanym punkcie, wartość liczbową $v = ds/dt$ i może być przedstawiony we współrzędnych prostokątnych za pomocą wielkości $v_x = dx/dt$, $v_y = dy/dt$, $v_z = dz/dt$. Rozpatrzony dalej przykład wyjaśni praktyczne znaczenie tych wyników (por. p. 1.4.4).

1.4.3. *Badanie przyspieszenia w ruchu krzywoliniowym.* W ruchu krzywoliniowym jednostajnym długości wektorów prędkości w różnych punktach toru są jednakowe, kierunki ich jednak ciągle się zmieniają, czyli nawet w ruchu jednostajnym krzywoliniowym wektor prędkości nie jest stały. A zatem w każdym ruchu krzywoliniowym (jednostajnym i niejednostajnym) można wyznaczyć wektorowy przyrost prędkości, a co za tym idzie i przyspieszenie. W tym celu od wektora prędkości $\mathbf{v}_2$ (rys. 1.11) w chwili



Rys. 1.11

późniejszej t_2 trzeba odjąć wektor prędkości $\mathbf{v}_1$ w chwili wcześniejszej t_1. Podczas odejmowania oba wektory muszą być odłożone z jednego punktu, trzeba zatem jeden z nich, np. $\mathbf{v}_2$, przenieść, zachowując jego długość i równoległość kierunku tak, by jego początek znalazł się w punkcie A. Wektor $\overrightarrow{CD}$ jest wektorowym przyrostem prędkości $\Delta\mathbf{v}$, odpowiadającym odstępowi czasu $\Delta t = t_2 - t_1$. Dzieląc ten przyrost prędkości przez czas Δt i przechodząc do granicy, gdy $\Delta t \to 0$, czyli obliczając pochodną $d\mathbf{v}/dt$, wyznaczamy *przyspieszenie wektorowe.* Temu przejściu do granicy towarzyszy też zmiana kierunku przyspieszenia średniego na kierunek przyspieszenia chwilowego. *Wektor przyspieszenia chwilowego różni się od wektora przyspieszenia średniego zarówno wartością, jak i kierunkiem.* Łatwo stwierdzić, że przyspieszenie chwilowe w ruchu krzywoliniowym skierowane jest zawsze w stronę wklęsłości toru, z wyjątkiem przypadku, gdy jest rozważane w punkcie przegięcia toru. Wtedy ma kierunek styczny do toru.

W układzie współrzędnych prostokątnych (x, y) można znaleźć wektor przyspieszenia w ruchu płaskim biorąc pierwszą pochodną prędkości $\mathbf{v}$ lub drugą pochodną promienia wodzącego $\mathbf{r}$ względem czasu:

$$\mathbf{a} = \frac{dv_x}{dt}\mathbf{i} + \frac{dv_y}{dt}\mathbf{j}$$

lub

$$\mathbf{a} = \frac{d^2x}{dt^2}\mathbf{i} + \frac{d^2y}{dt^2}\mathbf{j}.$$

Zastępując odpowiednio

$$\frac{dv_x}{dt} = \frac{d^2x}{dt^2} = a_x, \quad \frac{dv_y}{dt} = \frac{d^2y}{dt^2} = a_y$$

otrzymujemy

$$\mathbf{a} = a_x\mathbf{i} + a_y\mathbf{j}.$$

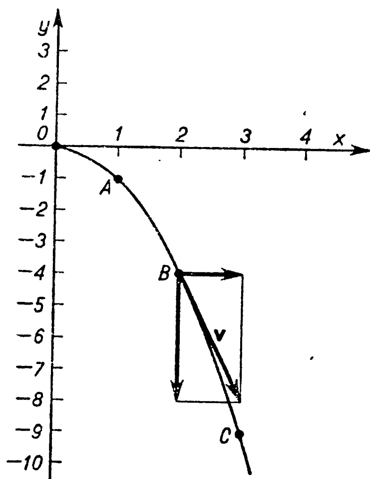
Uogólniając otrzymane wyrażenie na ruch przestrzenny w układzie (x, y, z) napiszemy

$$\mathbf{a} = a_x\mathbf{i} + a_y\mathbf{j} + a_z\mathbf{k},$$

gdzie

$$a_z = \frac{d^2z}{dt^2} = \frac{dv_z}{dt}.$$

1.4.4. *Przykład analizy ruchu krzywoliniowego.* Zastosujmy zdobyte wiadomości do przeanalizowania następującego przypadku. Niech w układzie współrzędnych (x, y) (rys. 1.12) punkt materialny odbywa ruch opisany zależnościami:



Rys. 1.12

$$x = v_0 t, \tag{1.18}$$

$$y = -\frac{gt^2}{2}, \tag{1.19}$$

gdzie v_0 i g są wielkościami niezależnymi od czasu. Zadanie nasze sprowadza się do: 1) wyznaczenia równania toru, 2) sporządzenia wykresu toru, 3) obliczenia prędkości i 4) obliczenia przyspieszenia.

1. Z charakteru równań wynika, że punkt materialny rozpoczyna ruch z początku układu współrzędnych (uzasadnić). Znak minus w równaniu (1.19) wskazuje, że ruch opisany tym równaniem odbywa się w kierunku ujemnym osi y. Postępując podobnie jak w przykładzie omówionym w p. 1.4.1 znajdujemy z (1.18)

$$t = \frac{x}{v_0},$$

a po podstawieniu do (1.19) otrzymujemy równanie toru badanego ruchu w postaci

$$y = -\frac{g}{2v_0^2}x^2. \tag{1.20}$$

2. Wykresem toru jest parabola przedstawiona na rys. 1.12. Jako jednostki na osiach x i y przyjęto wartości dróg przebytych w pierwszej jednostce czasu w odpowiednich ruchach, czyli wartość $v_0 \cdot 1$ na osi x oraz $g \cdot 1^2/2$ na osi y.

Punkty A, B, C są punktami toru mijanymi przez punkt materialny odpowiednio po 1, 2, 3 sekundach.

3. Prędkość w badanym ruchu po czasie t znajdziemy obliczając

$$v_x = \frac{dx}{dt} = v_0 \quad \text{(prędkość stała, a zatem ruch jednostajny wzdłuż osi } x\text{)},$$

$v_y = \frac{dy}{dt} = -gt$ (prędkość proporcjonalna do czasu, skierowana przeciwnie niż oś y, a zatem ruch jednostajnie przyspieszony w kierunku ujemnym osi y).

Wypadkowa prędkość ruchu badanego równa się

$$v = \sqrt{v_x^2+v_y^2} = \sqrt{v_0^2+g^2t^2}.$$

4. Obliczamy przyspieszenie

$$a_x = \frac{dv_x}{dt} = 0, \quad a_y = \frac{dv_y}{dt} = -g.$$

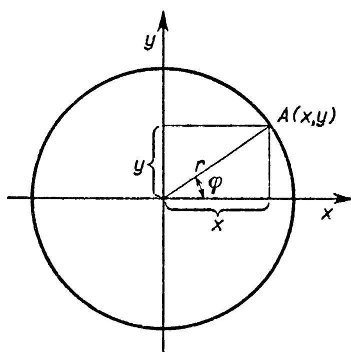
Niech kierunek osi x będzie poziomy, osi y — pionowy, a wartość g niech odpowiada przyspieszeniu ziemskiemu. Wtedy omówiony przykład będzie odpowiadał tzw. *rzutowi poziomemu* (por. § 4.7).

Rysunek 1.12 uzupełniamy przedstawiając graficznie prędkość np. w punkcie B toru, tzn. po 2 sekundach od chwili rozpoczęcia ruchu. Prędkość v_x równa się v_0, prędkość $v_y = -2g$. Dodając geometrycznie odpowiednie wektory otrzymujemy prędkość wypadkową w punkcie B, skierowaną stycznie do paraboli (warto to samo wykonać np. dla punktów A i C toru).

Zadanie. Kula wyrzucona poziomo z prędkością v_0 obniżyła się o h metrów względem poziomu początkowego. O jaką odległość x przesunęła się ona w tym samym czasie? Jaką miała prędkość końcową?

1.5. Ruch po okręgu

W rozważanym przypadku znając kształt toru wyznaczymy zależności $x = x(t)$ oraz $y = y(t)$. Jak widać z rys. 1.13, gdy początek układu współrzędnych (x, y) jest



Rys. 1.13

tak dobrany, że leży w środku koła, to współrzędne (x, y) punktu obiegającego okrąg można przedstawić jako

$$x = r\cos\varphi, \tag{1.21}$$

$$y = r\sin\varphi, \tag{1.22}$$

gdzie r — promień koła — ma wartość stałą, a kąt φ zakreślany przez promień wodzą-

cy jest funkcją czasu, czyli $\varphi = \varphi(t)$. Kąt φ można traktować jako tzw. *drogę kątową* promienia wodzącego. Charakter zależności drogi kątowej od czasu decyduje o tym, czy ruch po okręgu jest jednostajny, czy zmienny.

Obliczmy składowe prędkości liniowej $\mathbf{v}$ równe dx/dt i dy/dt różniczkując wyrażenia (1.21) i (1.22):

$$v_x = \frac{dx}{dt} = -r\frac{d\varphi}{dt}\sin\varphi,$$

$$v_y = \frac{dy}{dt} = r\frac{d\varphi}{dt}\cos\varphi.$$

Wektor prędkości można przedstawić jako sumę wektorową

$$\mathbf{v} = v_x\mathbf{i}+v_y\mathbf{j}.$$

Wartość liczbowa wektora $\mathbf{v}$ w ruchu po okręgu wynosi

$$v = \sqrt{v_x^2+v_y^2} = \sqrt{r^2\left(\frac{d\varphi}{dt}\right)^2(\sin^2\varphi+\cos^2\varphi)},$$

czyli

$$v = r\frac{d\varphi}{dt}.$$

Pochodną drogi kątowej względem czasu $d\varphi/dt$ nazywamy *prędkością kątową* i oznaczamy symbolem ω. A zatem

$$v = \omega r. \tag{1.23}$$

Prędkość kątowa w układzie SI wyrażana jest w radianach na sekundę (skrót rad/s).

Warto podkreślić, że prędkość kątowa jest podstawową wielkością opisującą ruch punktu po okręgu. Jest ona związana wzorem (1.23) z wartością liczbową prędkości v, którą dla podkreślenia jej odrębności nazywa się czasem *prędkością liniową.*

Prędkość kątową umówiono się traktować jako wektor prostopadły do płaszczyzny toru kołowego, wyprowadzony z jego środka. Kierunek tego wektora określa reguła korkociągu. Zgodnie z tą regułą, jeśli rączkę korkociągu ustawionego prostopadle do płaszczyzny toru kołowego obracać zgodnie z kierunkiem ruchu punktu po okręgu, to kierunek przesuwania się korkociągu określi kierunek wektora prędkości kątowej $\boldsymbol{\omega}$. Łatwo sprawdzić, że po przypisaniu prędkości kątowej $\boldsymbol{\omega}$ charakteru wektorowego można wzór (1.23) przepisać w postaci

$$\mathbf{v} = \boldsymbol{\omega}\times\mathbf{r}. \tag{1.24}$$

1.5.1. *Ruch jednostajny po okręgu.* W tym ruchu prędkość kątowa ma wartość stałą $\omega = \text{const}$, a zatem droga kątowa φ rośnie proporcjonalnie do czasu:

$$\varphi = \omega t.$$

Tym razem równania ruchu mają postać

$$x = r\cos\omega t, \quad y = r\sin\omega t,$$

a składowe prędkości wynoszą odpowiednio

$$v_x = -\omega r \sin \omega t, \tag{1.25a}$$

$$v_y = \omega r \cos \omega t. \tag{1.25b}$$

Wartość liczbowa prędkości wyraża się oczywiście nadal wzorem (1.23), z tym że w tym wzorze ω ma obecnie wartość stałą, równą $2\pi/T$, gdzie T — *okres*, tj. czas jednego pełnego obiegu punktu materialnego po okręgu.

Ze stałości prędkości kątowej wynika stałość prędkości liniowej

$$v = \omega r = \frac{2\pi r}{T}.$$

Jak już stwierdziliśmy poprzednio, w każdym ruchu krzywoliniowym (nawet jednostajnym) istnieje przyspieszenie. Obliczmy składowe przyspieszenia a_x i a_y w ruchu jednostajnym po okręgu:

$$a_x = \frac{d^2x}{dt^2} = -\omega^2 r \cos \omega t, \tag{1.26}$$

$$a_y = \frac{d^2y}{dt^2} = -\omega^2 r \sin \omega t. \tag{1.27}$$

Uwzględniając, że

$$\mathbf{a} = a_x \mathbf{i} + a_y \mathbf{j}$$

oraz

$$a = \sqrt{a_x^2 + a_y^2},$$

po podstawieniu wyrażeń na a_x i a_y otrzymuje się

$$a = \omega^2 r. \tag{1.28}$$

Łatwo sprawdzić, że w zapisie wektorowym wzór na przyspieszenie przyjmuje postać

$$\mathbf{a} = -\omega^2 \mathbf{r}. \tag{1.29}$$

Znak minus w tym wzorze wiąże się z kierunkiem dośrodkowym przyspieszenia. Ta cecha przyspieszenia w ruchu jednostajnym po okręgu zdecydowała o nadaniu mu nazwy *przyspieszenia dośrodkowego*.

Pytania. 1. Przyspieszenie dośrodkowe nazywa się niekiedy *przyspieszeniem normalnym*, podkreślając w ten sposób jego prostopadłość do innego wektora charakteryzującego ruch po okręgu. Jaki to jest wektor?

2. Wykazać, że w ruchu jednostajnym po okręgu składowa przyspieszenia styczna do toru równa się zeru.

Uwzględniając (1.23) można przyspieszenie dośrodkowe wyrazić za pomocą prędkości liniowej i promienia, a mianowicie

$$a = \frac{v^2}{r}. \tag{1.30}$$

Trzeba raz jeszcze podkreślić, że *w ruchu jednostajnym po okręgu, mimo istnienia przyspieszenia dośrodkowego, wartość liczbowa prędkości liniowej v nie ulega zmianie.* Istnienie tego przyspieszenia wpływa jedynie na zakrzywienie toru, czyli na zmianę kierunku wektora $\mathbf{v}$.

W ruchu jednostajnym po okręgu wektor prędkości kątowej $\boldsymbol{\omega}$ jest stały zarówno co do wartości, jak i co do kierunku. A zatem w tym ruchu *przyspieszenie kątowe* α, zdefiniowane jako pierwsza pochodna prędkości kątowej ω względem czasu lub druga pochodna drogi kątowej φ względem czasu, równa się zeru:

$$\alpha = \frac{d\omega}{dt} = \frac{d^2\varphi}{dt^2} = 0.$$

1.5.2. *Ruch niejednostajny po okręgu.* Inaczej przedstawia się sprawa w ruchu niejednostajnym po okręgu, tzn. gdy $\omega \neq$ const. Tym razem przyspieszenie kątowe α, zdefiniowane jak poprzednio, nie równa się zeru:

$$\alpha = \frac{d\omega}{dt} = \frac{d^2\varphi}{dt^2} \neq 0.$$

Jednostką przyspieszenia kątowego w układzie SI jest rad/s^2.

Według umowy przyspieszenie kątowe również jest traktowane jako wektor prostopadły do płaszczyzny toru kołowego wyprowadzony z jego środka. Kierunek wektora $\boldsymbol{\alpha}$ jest więc zawsze zgodny z kierunkiem $\boldsymbol{\omega}$, natomiast zwroty obu wektorów są zgodne w ruchu przyspieszonym, a przeciwne w ruchu opóźnionym.

Zadanie. Wykazać, że w przypadku ruchu niejednostajnego po okręgu (tzn. gdy $\omega \neq$ const, lecz jest pewną funkcją czasu) składowe przyspieszenia a_x i a_y wyrażają się następującymi wzorami:

$$a_x = \frac{\alpha}{\omega} v_x - \omega^2 x, \quad a_y = \frac{\alpha}{\omega} v_y - \omega^2 y.$$

Podane wyżej zależności wykorzystamy do określenia wektora $\mathbf{a}$ w ruchu niejednostajnym po okręgu. Sumując wektorowo otrzymujemy

$$\mathbf{a} = a_x \mathbf{i} + a_y \mathbf{j},$$

czyli

$$\mathbf{a} = \frac{\alpha}{\omega} \mathbf{v} - \omega^2 \mathbf{r}. \tag{1.31}$$

Wynik odczytamy w ten sposób, że wektor $\mathbf{a}$ jest sumą wektorową dwóch przyspieszeń, a mianowicie $(\alpha/\omega)\mathbf{v}$ i $-\omega^2\mathbf{r}$. Wektor napisany na drugim miejscu jest znanym już nam przyspieszeniem dośrodkowym (normalnym), występującym także w ruchu jednostajnym po okręgu, odpowiedzialnym za zakrzywienie toru

$$\mathbf{a}_n = -\omega^2 \mathbf{r}.$$

Drugi wektor składowy o kierunku zgodnym z wektorem $\mathbf{v}$ (a więc styczny do okręgu) nazywać będziemy *przyspieszeniem stycznym* i oznaczać $\mathbf{a}_t$

$$\mathbf{a}_t = \frac{\alpha}{\omega} \mathbf{v}. \tag{1.32}$$

Przyspieszenie styczne odpowiedzialne jest za zmiany wartości liczbowej $\mathbf{v}$.

Zależność (1.31) można więc przedstawić jako

$$\mathbf{a} = \mathbf{a}_n + \mathbf{a}_t.$$

Jak widać z przekształcenia wzoru (1.32), wartość liczbowa wektora $\mathbf{a}_t$ równa się

$$a_t = \alpha r. \tag{1.33}$$

Pytania. 1. Jakim ruchom odpowiadają następujące zestawienia wartości a_n i a_t:

a) $a_n = 0$, $a_t = 0$ b) $a_n \neq 0$, $a_t = 0$ c) $a_n = 0$, $a_t \neq 0$ d) $a_n \neq 0$, $a_t \neq 0$

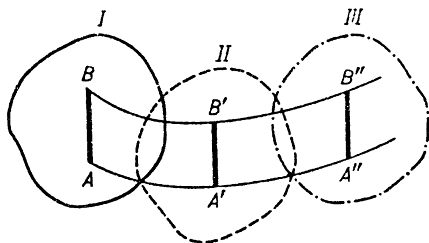
2. Jakim ruchom po okręgu odpowiadają następujące charakterystyki:

a) $\omega = \text{const}$, $\alpha = 0$ b) $\omega \neq \text{const}$, $\alpha = \text{const}$ c) $\omega \neq \text{const}$, $\alpha \neq \text{const}$

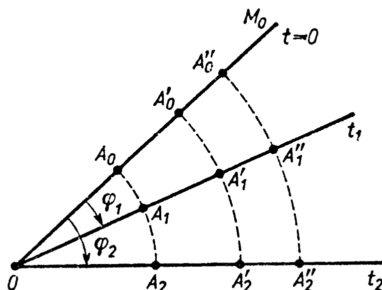
1.6. Kinematyka bryły sztywnej

Dotychczas zajmowaliśmy się ruchem punktu materialnego. Obecnie przechodzimy do zagadnień nieco bardziej skomplikowanych, a mianowicie zajmiemy się kinematyką bryły sztywnej.

Przez *bryłę sztywną* rozumiemy ciało, które pod działaniem dowolnie wielkich sił nie ulega ani odkształceniu postaci (zmiana kształtu), ani odkształceniu objętości. Odległość dwóch dowolnych punktów bryły sztywnej pozostaje niezmienna.



Rys. 1.14



Rys. 1.15

Bryła sztywna może wykonywać dwa rodzaje ruchów prostych: ruch postępowy i ruch obrotowy. (Jeśli bryła odbywa równocześnie oba te rodzaje ruchów, to powstaje ruch złożony, którym bliżej nie będziemy się zajmowali.) *Ruch postępowy bryły sztywnej* jest to taki ruch, przy którym dowolny odcinek łączący dwa punkty bryły, np. A i B (rys. 1.14), zachowuje stale położenie do siebie równoległe. Wszystkie punkty bryły sztywnej, odbywającej ruch postępowy, zakreślają drogi równe oraz mają jednakowe prędkości i przyspieszenia. W związku z tym badanie ruchu postępowego bryły sztywnej sprowadza się do badania ruchu jakiegokolwiek dowolnie wybranego punktu bryły.

Jeśli bryła sztywna wprawiona jest w *ruch obrotowy*, można w niej wyodrębnić szereg punktów nie poruszających się. Zbiór tych punktów leżących na jednej prostej stanowi tzw. *oś obrotu*. Oś obrotu jest stała, jeżeli z biegiem czasu nie zmienia swego położenia ani w ciele ani w przestrzeni. Pozostałe punkty bryły zataczają tory kołowe w płaszczyznach prostopadłych do osi. Promienie tych kół równe są odległościom rozpatrywanych punktów od osi obrotu. Zajmiemy się dalej ruchem obrotowym bryły sztywnej dokoła stałej osi.

Niech płaszczyzna rys. 1.15 przedstawia płaszczyznę jednego z możliwych przekrojów bryły prostopadłych do osi obrotu, a punkt O niech będzie śladem osi obrotu.

Rozważmy położenie dowolnie wybranego w tej płaszczyźnie promienia wodzącego OM_0 w chwilach $t = 0$, t_1 i t_2. Punkty A_0, A_0' i A_0'' po czasie t_1 zajmą położenia A_1, A_1' i A_1'', a po czasie t_2 — położenia A_2, A_2' i A_2'' Każdy z tych punktów w odstępie czasu $t_2—t_1$ zakreślił drogę równą odpowiednio łukowi A_1A_2, $A_1'A_2'$ i $A_1''A_2''$. Długości tych łuków są różne, a więc i prędkości liniowe poszczególnych punktów są niejednakowe, są zależne od ich odległości od osi obrotu.

W związku z zależnością prędkości liniowej od odległości od osi obrotu wygodniej jest charakteryzować ruch obrotowy za pomocą innej wielkości, a mianowicie *prędkości kątowej* ω. Niech A_0O oznacza położenie początkowe w chwili $t = 0$ dowolnego promienia wodzącego pomyślanego w bryle sztywnej. Po czasie t_1 ten promień wodzący zajmuje nowe położenie A_1O przesunięte względem poprzedniego o kąt φ_1, po czasie t_2 — położenie A_2O przesunięte względem początkowego o kąt φ_2. A zatem w czasie $\Delta t = t_2—t_1$ promień wodzący zakreślił kąt $A_1OA_2 = \varphi_2—\varphi_1 = \Delta\varphi$: Średnio na jednostkę czasu przypada kąt obrotu

$$\bar{\omega} = \frac{\Delta\varphi}{\Delta t}.$$

Wielkość $\bar{\omega}$ nazywamy *średnią prędkością kątową.*

Jeśli ruch obrotowy jest jednostajny, to znaleziona prędkość kątowa średnia jest zarazem prędkością kątową chwilową. Ogólnie jednak prędkość kątową (chwilową) wyznaczamy przez przejście do granicy, gdy $\Delta t \to 0$. A zatem

$$\omega = \frac{d\varphi}{dt}. \tag{1.34}$$

Z rozważań dotyczących ruchu po okręgu wiemy już, że ω w układzie SI wyraża się w rad/s.

Ze względu na sztywność obracającej się bryły zachowanie się każdego dowolnie wybranego promienia wodzącego w każdym przekroju bryły prostopadłym do osi obrotu będzie charakteryzowała ta sama prędkość kątowa ω. Można zatem tę wielkość traktować jako podstawową wielkość kinematyczną charakteryzującą ruch obrotowy. Ruch obrotowy bryły sztywnej jest jednostajny, gdy $\omega = \text{const}$.

Prędkość kątową umówiono się traktować jako wektor skierowany wzdłuż osi obrotu bryły, o punkcie przyłożenia wybranym dowolnie na tej osi (wektor przesuwalny — por. Wiadomości wstępne, § 6). Zwrot tego wektora jest określony regułą korkociągu. Ustawiamy korkociąg wzdłuż osi obrotu i rączkę korkociągu obracamy w kierunku obrotu bryły. Kierunek przesuwania się korkociągu wyznacza zwrot wektora prędkości kątowej.

W ruchu obrotowym niejednostajnym prędkość kątowa nie jest stała. Możemy więc posługiwać się pojęciem *przyspieszenia kątowego* α, zdefiniowanego jak poprzednio:

$$\alpha = \frac{d\omega}{dt}. \tag{1.35}$$

Przyspieszenie kątowe wyraża się w rad/s^2.

Znów ze względu na sztywność bryły przyspieszenie kątowe α charakteryzuje zachowanie się każdego z jej punktów (oczywiście nie leżących na osi obrotu).

Przyspieszenie kątowe jest również wektorem. W przypadku ruchu obrotowego dokoła stałej osi jest ono skierowane wzdłuż osi obrotu. Zwrot wektora przyspieszenia kątowego jest zgodny ze zwrotem wektora prędkości kątowej w ruchu przyspieszonym, przeciwny — w ruchu opóźnionym.

Dla każdego z punktów bryły (nie leżących na osi obrotu) są ponadto spełnione zależności:

$$v = \omega r, \qquad a_t = \alpha r,$$

gdyż każdy z tych punktów odbywa ruch po okręgu o promieniu r.

Pytania i zadania

1. Podać definicję podstawowych pojęć kinematycznych jak: a) ruch, b) układ odniesienia, c) tor, d) droga.

2. Podać definicje wielkości kinematycznych: a) prędkości i b) przyspieszenia.

3. Sklasyfikować ruch punktu materialnego z punktu widzenia: a) toru, b) prędkości, c) przyspieszenia.

4. Jak się przedstawia wykres zależności: a) drogi i prędkości od czasu w ruchu jednostajnym, b) drogi, prędkości i przyspieszenia od czasu w ruchu jednostajnie zmiennym.

5. Scharakteryzować metodę badania ruchu krzywoliniowego płaskiego.

6. Zdefiniować prędkość $\mathbf{v}$ w ruchu krzywoliniowym.

7. Opisać metodę wyznaczania przyspieszenia średniego. Jak się odbywa przejście od przyspieszenia średniego do przyspieszenia chwilowego w ruchu krzywoliniowym?

8. Jak można wyrazić wartość liczbową wektora $\mathbf{a}$?

9. Na czym polega badanie ruchu krzywoliniowego na znanym torze, np. na okręgu?

10. Przeanalizować ruch jednostajny po okręgu z punktu widzenia: a) prędkości $\mathbf{v}$ (wartość liczbowa i kierunek), b) prędkości kątowej ω, c) składowych prędkości $\mathbf{v}_x$ i $\mathbf{v}_y$, d) składowych przyspieszenia $\mathbf{a}_x$ i $\mathbf{a}_y$, c) przyspieszenia $\mathbf{a}$ (wartość liczbowa i kierunek).

11. W analogiczny sposób jak w zad. 10 przeanalizować ruch niejednostajny po okręgu.

12. Jakie zmiany w ruchu niejednostajnym po okręgu wywołuje: a) przyspieszenie normalne i b) przyspieszenie styczne?

13. Sformułować definicje: a) bryły sztywnej, b) stałej osi obrotu bryły.

14. Jakie wielkości charakteryzują ruch bryły sztywnej: a) obrotowy jednostajny, b) obrotowy jednostajnie zmienny?

15. Pociąg jadący z prędkością 60 km/h zaczyna hamować w odległości 1 km od stacji. Zakładając, że ruch jest jednostajnie opóźniony obliczyć opóźnienie oraz czas hamowania.

Odp. 5/36 m/s², 120 s.

16. Samochód rusza z miejsca z przyspieszeniem $a = 2$ m/s². Jaką drogę przejedzie w ciągu dziesiątej sekundy ruchu i jaką prędkość będzie miał po dziesięciu sekundach?

Odp. 19 m, 20 m/s.

17. Na pewnym odcinku swego toru o długości $s = 800$ m pociąg jedzie wzdłuż łuku koła o promieniu $r = 600$ m ruchem jednostajnie opóźnionym. Prędkość pociągu na początku tego odcinka wynosi $v_0 = 54$ km/h, a na końcu $v = 18$ km/h. Obliczyć czas przejazdu tego odcinka oraz całkowite przyspieszenie pociągu na początku i na końcu tego łuku.

Odp. 80 s, 0,395 m/s², 0,131 m/s².

18. Obliczyć prędkość i przyspieszenie po upływie t sekund w ruchu opisanym zależnością (1.9) (jest to tzw. ruch harmoniczny).

ROZDZIAŁ 2

Dynamika ruchu postępowego

2.1. Pierwsza zasada dynamiki

Opisując różne rodzaje ruchu z punktu widzenia kinematyki podawaliśmy ich formalną charakterystykę. W dynamice interesują nas warunki, w jakich poszczególne ruchy powstają, a przede wszystkim przyczyny ich powstawania. Rozpoczniemy od rozważań dynamicznych związanych z ruchem postępowym brył, które będziemy traktowali jako punkty materialne. Dynamikę układu punktów materialnych omówimy w § 2.7 i 2.8. Dynamice ruchu obrotowego brył poświęcimy rozdz. 5.

Ustalenie zasad dynamiki było równoznaczne z obaleniem fałszywych poglądów panujących od czasów starożytnych a dotyczących przyczyn powstawania różnych rodzajów ruchu. Do obalenia tych poglądów przyczyniło się wprowadzenie do nauki metody doświadczalnej przez Galileusza w XVI wieku. Badania zapoczątkowane przez Galileusza podjęte zostały następnie przez Newtona, któremu zawdzięczamy ustalenie podstaw dynamiki.

Zasady dynamiki podane zostały przez Newtona jako tzw. prawa ruchu. Zgodnie z pierwszą zasadą dynamiki *ciało nie poddane działaniu żadnej siły albo poddane działaniu sił równoważących się pozostaje w spoczynku lub porusza się ruchem jednostajnym prostoliniowym.*

Słuszność pierwszej części tej zasady, odnosząca się do przypadku, gdy na ciało nie działa żadna siła, nie może być na Ziemi doświadczalnie sprawdzona, nie możemy bowiem stworzyć na Ziemi takich warunków, aby ciało było wolne od działania sił. Druga część nadaje się do doświadczalnego sprawdzenia. Po dokładnym zrównoważeniu sił oporu przez siłę ciągnącą ciało, mające pewną prędkość **v**, zachowa tę prędkość niezmienną zarówno co do wartości, jak i co do kierunku, tzn. poruszać się będzie ruchem jednostajnym prostoliniowym.

Pierwsza zasada dynamiki nosi nazwę *zasady bezwładności*. Przez *bezwładność* rozumiemy właściwość ciała decydującą o tym, że ciało bez działania sił nie może zmienić ani wartości, ani kierunku swej prędkości.

2.2. Druga zasada dynamiki

Wszelkie zmiany prędkości mogą zachodzić jedynie pod działaniem sił. Musi więc istnieć związek między siłą a zmianami prędkości. Ta zależność jest treścią drugiej

zasady dynamiki, którą można ująć następująco: *siła jest proporcjonalna do przyspieszenia, które wywołuje*:

$$\mathbf{F} = m\mathbf{a}, \tag{2.1}$$

gdzie **F** oznacza siłę, m — masę, **a** — przyspieszenie.

Równanie (2.1) jest równaniem wektorowym: wektor przyspieszenia ma kierunek i zwrot zgodny z kierunkiem i zwrotem działającej siły. Masa m ciała jest miarą jego bezwładności*.

Określenie jednostki siły w układzie SI wynika z równania (2.1). Jednostką siły w układzie SI jest taka siła, która działając na ciało o masie 1 kg nadaje mu przyspieszenie równe 1 m/s^2. Siłę tę nazwano *niutonem* (skrót N)**.

$$1\,\mathrm{N} = 1\frac{\mathrm{kg}\cdot\mathrm{m}}{\mathrm{s}^2}.$$

Wracając do (2.1) trzeba podkreślić, że o proporcjonalności siły do przyspieszenia mówimy w przypadku oddziaływania różnych sił na tę samą masę m ($m = \mathrm{const}$). Jeżeli natomiast tą samą siłą F ($F = \mathrm{const}$) działać będziemy kolejno na ciała o różnych masach $m_1, m_2 \ldots$, to obowiązywać będą zależności

$$F = m_1 a_1 = m_2 a_2 = \ldots,$$

skąd

$$\frac{m_1}{m_2} = \frac{a_2}{a_1}.$$

A zatem przyspieszenia uzyskane przez różne ciała pod działaniem tej samej siły są odwrotnie proporcjonalne do mas tych ciał.

Wyrażenie $\mathbf{F} = m\mathbf{a}$ jest również słuszne, gdy ciało o masie m poddane jest jednoczesnemu działaniu kilku sił: $\mathbf{F}_1$, $\mathbf{F}_2$ itd. W tym przypadku **F** jest sumą geometryczną (wypadkową) wszystkich sił działających, a **a** jest przyspieszeniem ciała.

Z wyrażeniem (2.1) wiąże się ściśle tzw. *dynamiczny pomiar siły*. Wystarczy znać masę ciała i wyznaczyć uzyskane przez nią przyspieszenie, aby móc obliczyć wartość działającej siły***.

2.3. Ogólniejsze ujęcie drugiej zasady dynamiki

Do wyrażenia (2.1) podstawiamy

$$\mathbf{a} = \frac{\mathbf{v}_2 - \mathbf{v}_1}{t_2 - t_1}$$

* Przypominamy o konieczności rozróżniania pojęcia masy (miary bezwładności) i ciężaru ciała (siły przyciągania grawitacyjnego) (por. § 4.4).

** Ta sama jednostka jest jednostką siły w układzie MKS. W układzie CGS jednostką siły jest dyna. *Dyna* jest to siła, która ciału o masie 1 g nadaje przyspieszenie 1 cm/s^2. Łatwo obliczyć, że 1 N = = 10^5 dyn.

*** W praktyce szerzej stosowany jest *pomiar statyczny siły*, który polega na wywoływaniu odkształceń w ciałach sprężystych i porównywaniu ich z odkształceniami wywołanymi przez siły znane (np. posługiwanie się wycechowanym dynamometrem). Opieramy się przy tym na prostej proporcjonalności sił i wywołanych przez nie odkształceń sprężystych (por. § 8.2).

Jeżeli badaną siłą jest ciężar ciała, to do jej wyznaczania stosuje się różnego typu wagi (por. § 6.6).

i otrzymujemy

$$\mathbf{F}(t_2-t_1) = m\mathbf{v}_2 - m\mathbf{v}_1.$$

Iloczyn siły i czasu jej działania nazywamy *popędem siły*. Jest to wektor o kierunku zgodnym z kierunkiem wektora **F**.

Iloczyn masy i prędkości nosi nazwę *pędu*. Jest to również wektor. Kierunek wektora pędu **p** jest zgodny z kierunkiem prędkości **v**. Równanie ostatnie wyraża, że *wektor popędu siły jest równy wektorowemu przyrostowi pędu wywołanemu przez tę siłę*

$$\mathbf{F}(t_2-t_1) = \Delta\mathbf{p}, \tag{2.2}$$

czyli

$$\mathbf{F} = \frac{\Delta\mathbf{p}}{\Delta t}.$$

Jeśli w czasie t_2-t_1 wektor siły ulega zmianie, to otrzymane wyrażenie przedstawia siłę średnią w czasie Δt. Zakładając, że czas Δt zmierza do zera, znajdujemy *siłę chwilową w chwili t* jako pochodną pędu względem czasu:

$$\mathbf{F} = \frac{d(m\mathbf{v})}{dt} = \frac{d\mathbf{p}}{dt}. \tag{2.3}$$

Równanie (2.3) wyrażające siłę jako pochodną pędu względem czasu jest ogólniejsze od równania (2.1), to ostatnie bowiem jest słuszne jedynie wtedy, gdy ciała poruszają się z prędkościami małymi w porównaniu z prędkością światła. Gdy prędkość ciała jest porównywalna z prędkością światła, należy uwzględniać zmienność masy podczas ruchu (por. teoria względności — rozdz. 7). Zmienność masy wynikająca z ruchu ciała jest określona równaniem Einsteina:

$$m = \frac{m_0}{\sqrt{1-v^2/c^2}},$$

gdzie m oznacza masę ciała będącego w ruchu, m_0 — masę tegoż ciała w spoczynku, c — prędkość światła w próżni (ok. 300 000 km/s).

Ze wzoru wynika, że nawet wtedy, gdy prędkość ciała równa się 30 000 km/s, zmiana masy jest niewielka, mniejsza od 1%. Gdy prędkości zbliżają się do prędkości światła (co może występować np. w przypadku mikrocząstek), masa coraz szybciej rośnie. W tych warunkach wartość siły można obliczać tylko z równania (2.3): zamiast mechaniki niutonowskiej należy stosować mechanikę relatywistyczną.

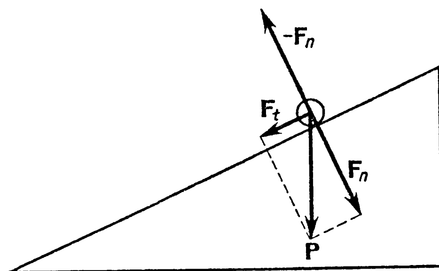
2.4. Trzecia zasada dynamiki

Trzecia zasada dynamiki, zwana również *zasadą akcji i reakcji*, dotyczy wzajemnego oddziaływania dwóch ciał (względnie układów ciał).

Jeżeli ciało A działa na ciało B siłą $\mathbf{F}_{AB}$, *to ciało B działa na ciało A siłą* $\mathbf{F}_{BA}$ *równą co do wartości, lecz przeciwnie skierowaną*:

$$\mathbf{F}_{AB} = -\mathbf{F}_{BA}.$$

Obie siły występują równocześnie, toteż nie można powiedzieć, która z nich jest siłą akcji, a która siłą reakcji, co widać wyraźnie np. w przypadku przyciągania grawitacyjnego dwóch ciał. Czasem jednak umownie odróżnia się siłę pierwotną — siłę akcji i siłę wtórną — siłę reakcji, np. w przypadku ciała spoczywającego na podstawie. Nacisk na podstawę traktuje się jako siłę akcji, a oddziaływanie podstawy jako siłę reakcji. Zestawienie tych sił w przypadku równi pochyłej przedstawia rys. 2.1. Siłę ciężkości **P** rozkładamy na dwie składowe: składową styczną $\mathbf{F}_t$ i składową normalną $\mathbf{F}_n$. Ta ostatnia jest właśnie siłą naciskającą równię. Odpowiada jej siła reakcji $-\mathbf{F}_n$. Druga składowa siły ciężkości, a mianowicie $\mathbf{F}_t$, jest siłą wprawiającą ciało w ruch po równi. Chcąc utrzymać to ciało w spoczynku, należy tę składową zrównoważyć dodatkową siłą, również styczną do równi, równą co do wartości $\mathbf{F}_t$, lecz przeciwnie skierowaną.



Rys. 2.1

Siły wzajemnego oddziaływania między ciałami występują zarówno wtedy, gdy ciała spoczywają, jak i wtedy, gdy ciała są w ruchu, np. przyciąganie grawitacyjne między Ziemią i Słońcem, między poszczególnymi planetami itd. Przyciąganie grawitacyjne występuje zarówno w próżni, jak też wtedy, gdy przestrzeń między przyciągającymi się ciałami wypełniona jest dowolnym ośrodkiem.

Często ośrodek pośredniczy w przekazywaniu działania od jednego ciała do drugiego. Na przykład do sztywnej belki przymocowana jest sprężyna, na której wisi ciało o ciężarze P. Działanie siły **P** przenosi się za pośrednictwem sprężyny na belkę.

2.5. Dynamika ruchu po okręgu

Z kolei zajmiemy się siłami występującymi podczas ruchu po okręgu.

2.5.1. *Siła dośrodkowa.* Rozważmy ruch jednostajny po okręgu z punktu widzenia dynamiki. Zgodnie z I zasadą dynamiki tylko ruch jednostajny prostoliniowy może istnieć bez działania sił. Ruch jednostajny po okręgu wymaga już istnienia siły. Według II zasady dynamiki wartość liczbowa tej siły wyraża się zależnością

$$F = ma.$$

Uwzględniając wyrażenie (1.29) i (1.30) otrzymujemy

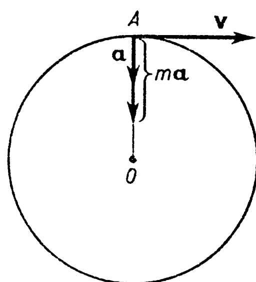
$$F = \frac{mv^2}{r} \quad \text{oraz} \quad F = m\omega^2 r. \tag{2.4}$$

Kierunek tej siły jest zgodny z kierunkiem przyspieszenia, a więc wzdłuż promienia do środka koła. Stąd pochodzi nazwa *siła dośrodkowa.*

Wstawiając zamiast ω wartość $2\pi/T$ otrzymujemy jeszcze inną postać zależności (2.4):

$$F = m\frac{4\pi^2}{T^2}r. \tag{2.5}$$

Rysunek 2.2 przypomina, że kierunek siły dośrodkowej jest prostopadły do kierunku prędkości. Jest to uwaga bardzo istotna. Jeśli kiedykolwiek przy badaniu ruchu ciała stwierdzimy, że podlega ono działaniu siły stałej co do wartości i skierowanej stale prostopadle do prędkości ruchu, to będziemy mogli twierdzić, że torem ruchu jest okrąg (np. ruch promieni katodowych w polu magnetycznym).



Rys. 2.2

Można wymienić wiele przykładów siły dośrodkowej. Gdy kamień przymocowany do sznurka wprawiamy w ruch po kole, to siłę dośrodkową wywiera nasza ręka za pośrednictwem napiętego sznurka. Gdy pociąg posuwa się po zakrzywionym torze, to sprężyste oddziaływanie zewnętrznej szyny stanowi siłę dośrodkową. Jeśli przyjmiemy, że Księżyc krąży dokoła Ziemi po torze kołowym, to siłę dośrodkową stanowi przyciąganie grawitacyjne Ziemi.

Podczas krążenia elektronu po kołowej orbicie dokoła jądra atomu siłę dośrodkową stanowi elektryczne przyciąganie ujemnie naładowanego elektronu przez dodatnio naładowane jądro atomowe.

Załóżmy, że w pewnej chwili przestaje działać na ciało siła dośrodkowa, np. przerywa się sznurek, do którego był przymocowany poruszający się kamień. Zgodnie z zasadą bezwładności ruch ciała nie ustaje, lecz trwa dalej jako ruch jednostajny wzdłuż stycznej do toru kołowego. Oczywiście, sprawa się komplikuje, jeśli prócz siły dośrodkowej działają na ciało jeszcze inne siły. Grudki błota odlatują od koła rowerowego po torze krzywoliniowym, gdyż do ruchu po stycznej dodaje się ruch wywołany działaniem siły ciężkości. Podobnie przedstawia się ruch odlatujących od tarczy szlifierki opiłków szlifowanych na niej przedmiotów.

2.5.2. *Siła odśrodkowa reakcji.* Zgodnie z trzecią zasadą dynamiki działaniu siły dośrodkowej na ciało krążące po okręgu musi towarzyszyć działanie siły odśrodkowej na tzw. „więzy". Przez więzy rozumiemy te ciała, które wymuszają ruch po kole. W naszych przykładach takimi więzami będą: ręka wprawiająca kamień w ruch po kole za pośrednictwem sznurka, szyna kolejowa, Ziemia i jądro atomowe.

Siła odśrodkowa $\mathbf{F}_{\text{odśr}}$ jest równa co do wartości sile dośrodkowej $\mathbf{F}_{\text{dośr}}$, lecz ma zwrot przeciwny:

$$\mathbf{F}_{\text{dośr}} = -\mathbf{F}_{\text{odśr}}.$$

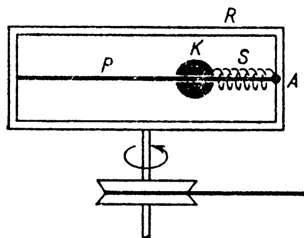
Siła dośrodkowa nie równoważy się z siłą odśrodkową, gdyż obie działają na różne ciała.

Pytanie. Jakie siły działają na punkt materialny, wykonujący ruch niejednostajny po okręgu? Podać wzory dla obu sił składowych.

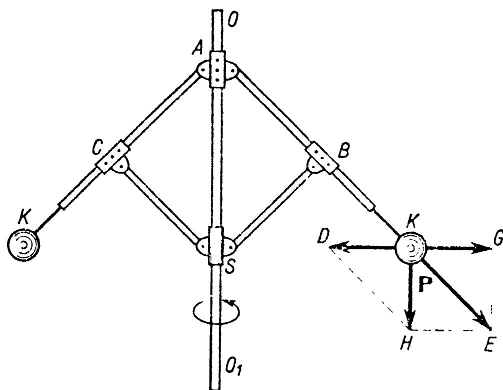
2.5.3. *Siła odśrodkowa bezwładności.* Przeanalizujemy obecnie ruch po okręgu z punktu widzenia obserwatora, który będąc związany z układem obracającym się (wykonując razem z nim ruch obrotowy) opisuje zachowanie się względem tego układu jakiegoś ciała ze swego otoczenia. Nazwijmy tego obserwatora obserwatorem ruchomym. Niech np. znajduje się on w pobliżu obwodu poziomej tarczy wykonującej obrót dokoła osi pionowej i trzyma w ręku kulę. Kula porusza się (razem z obserwatorem) po okręgu, a więc podlega działaniu siły dośrodkowej. Siłę tę wywiera ręka obserwatora. Ręka obserwatora stanowi w tym przykładzie więzy wymuszające ruch kuli po okręgu. Na rękę działa zatem siła odśrodkowa reakcji odpychająca rękę od osi obrotu na zewnątrz w kierunku zgodnym z promieniem okręgu.

Względem obserwatora kula jest w spoczynku. Nie obserwuje on jej ruchu po okręgu, odczuwa jednak działanie kuli na rękę. Logicznie rozumując obserwator uważałby, że na kulę działa pewna siła wzdłuż promienia na zewnątrz od osi obrotu i że kula przenosi to działanie na jego rękę. Zdaniem więc obserwatora ruchome ciało, spoczywające w układzie obracającym się (kula jest przecież nieruchoma względem obserwatora), podlega działaniu siły skierowanej wzdłuż promienia od środka. Tę siłę nazwano *siłą odśrodkową bezwładności* w odróżnieniu od siły odśrodkowej reakcji działającej na więzy. Do pojęcia siły odśrodkowej bezwładności wrócimy jeszcze w następnym paragrafie, gdzie udowodnimy, że co do wartości liczbowej siła odśrodkowa bezwładności jest równa sile dośrodkowej:

$$F_{\text{odśr bezwł}} = m\omega^2 r.$$



Rys. 2.3



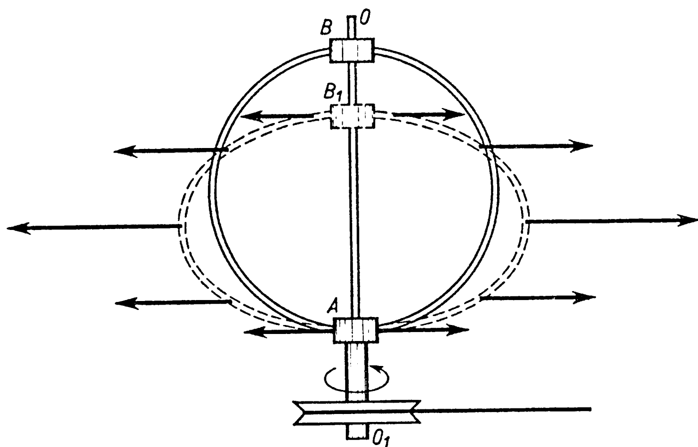
Rys. 2.4

Należy zwrócić uwagę na fakt, że według obserwatora nieruchomego z ruchem kołowym wiążą się tylko dwie siły: siła dośrodkowa, działająca na badane ciało, i siła odśrodkowa, działająca na więzy. Siła odśrodkowa bezwładności, zdaniem obserwatora nieruchomego, nie istnieje, dlatego też często jest nazywana *siłą odśrodkową pozorną.*

W wielu przypadkach wygodnie jest rozpatrywać zachowanie się ciał odbywających ruch kołowy z punktu widzenia obserwatora ruchomego. Rozpatrzmy kilka przykładów.

1. Na osi wirówki umieszczona jest pozioma ramka R (rys. 2.3), w której osadzona jest sprężyna S z przymocowaną do niej kulką K mogącą się przesuwać wzdłuż pręta P. Gdy wirówkę wprawimy w szybki ruch obrotowy, kulka przesunie się w stronę punktu A ściskając sprężynę. Można uważać (z punktu widzenia obserwatora ruchomego), że na kulkę działa siła odśrodkowa bezwładności, odsuwająca kulkę wzdłuż promienia od osi obrotu.

Obserwator spoza układu ruchomego, znający prawa rządzące ruchem po okręgu, powie, że kulka krąży po torze kołowym, gdyż podlega działaniu siły dośrodkowej ze strony ściśniętej sprężyny.



Rys. 2.5

2. Często wymienianym przykładem działania siły dośrodkowej jest siła występująca w ogólnie znanym regulatorze prędkości obrotowej, zwanym regulatorem Watta (rys. 2.4). Na osi pionowej OO_1 osadzone są dwa pręty AB i AC zakończone kulkami K. Podczas jednostajnego obrotu zarówno wysokość kulek, jak i suwaka S nie ulega zmianie. Na każdą z kulek działa siła ciężkości **P** pionowo w dół. Siłę tę można rozłożyć na dwie siły składowe: jedną $\overrightarrow{KE}$ działającą na przedłużeniu pręta, drugą $\overrightarrow{KD}$ — wzdłuż promienią okręgu zakreślanego przez kulkę K. Jedynym skutkiem działania siły $\overrightarrow{KE}$ jest naprężenie pręta. Druga składowa $\overrightarrow{KD}$ stanowi siłę dośrodkową potrzebną do utrzymania kulki w ruchu po okręgu.

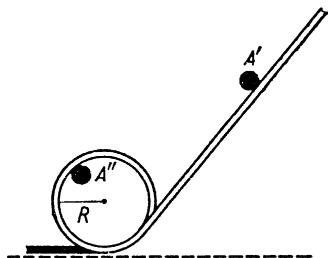
Inaczej przedstawia się rozumowanie z punktu widzenia obserwatora ruchomego. Według niego na kulkę działają dwie siły, a mianowicie siła odśrodkowa bezwładności $\overrightarrow{KG}$ (rys. 2.4) oraz siła ciężkości **P**. Kulka znajduje się na niezmiennej wysokości, gdy wypadkowa obu tych sił, tzn. siła $\overrightarrow{KE}$, jest skierowana wzdłuż przedłużenia osi pręta AK.

Jeżeli wzrasta prędkość ruchu obrotowego, to równocześnie rośnie siła odśrodkowa

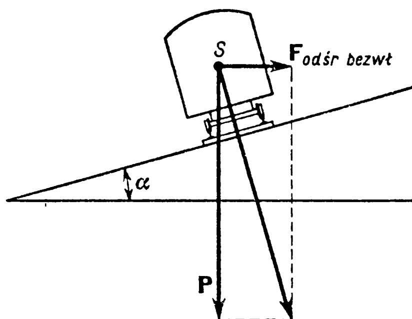
bezwładności, co powoduje zmianę kierunku siły wypadkowej. Ta nowa wypadkowa wywołuje zmianę położenia pręta AK, odchylając go o większy kąt od osi obrotu, a mianowicie doprowadzając do takiego położenia, przy którym kierunek nowej wypadkowej przypada znów na przedłużeniu pręta AK.

3. Spłaszczenie obręczy, występujące podczas ruchu obrotowego, tłumaczymy również działaniem siły odśrodkowej bezwładności (rys. 2.5).

Obręcz jest osadzona na osi OO_1 wirówki w ten sposób, że dolne zamocowanie A jest nieruchome, górne B — ruchome. Na cząstki obręczy leżące dalej od osi działa większa siła odśrodkowa bezwładności, gdyż zgodnie ze wzorem (2.4) przy stałym m i ω siła odśrodkowa jest proporcjonalna do r.



Rys. 2.6



Rys. 2.7

W podobny sposób wyjaśniamy również spłaszczenie Ziemi, wywołane jej ruchem obrotowym dokoła własnej osi, datujące się jeszcze z tego okresu, gdy Ziemia nie miała zastygniętej skorupy.

4. Działaniem siły odśrodkowej bezwładności tłumaczymy też działanie wirówek. Tym razem mamy cząstki o różnych masach, odbywające ruch obrotowy o tej samej prędkości kątowej. Siła odśrodkowa bezwładności działająca na cząstki równoodległe od osi obrotu jest proporcjonalna do ich masy. Cząstki o większej masie, podlegające większym siłom odśrodkowym bezwładności, oddalają się bardziej od osi obrotu i wykonują ruch po okręgach o większych promieniach. W ten sposób oddziela się np. cząstki tłuszczu (śmietankę) od mleka.

Do rozdzielania cząstek o mało różniących się masach służą tzw. ultrawirówki. W nowoczesnych ultrawirówkach liczba obrotów na sekundę sięga dwóch tysięcy.

P y t a n i a. Pojęcie siły odśrodkowej bezwładności pozwala tłumaczyć wiele na pozór nie zrozumiałych faktów. Zastanówmy się np. dlaczego przy odpowiedniej prędkości wiadra obracanego w płaszczyźnie pionowej nie wylewa się z niego woda? Dlaczego cyklista zjeżdżający w cyrku po ukośnym torze i wjeżdżający następnie na pętlę nie spada z niej w górnej jej części (uproszczony rys. 2.6), mimo że jedzie przez pewien czas „do góry nogami”? Dlaczego szyny kolejowe na zakrętach układane są w ten sposób, że płaszczyzna szyn jest nachylona pod pewnym kątem względem poziomu (rys. 2.7)? Dlaczego samochód jadący odcinkami drogi o powierzchni płaskiej, wklęsłej i wypukłej wywiera w każdym przypadku inny nacisk?

2.6. Zasada względności ruchów w mechanice klasycznej

We wstępie do kinematyki scharakteryzowaliśmy ruch ciała jako zmianę jego położenia w stosunku do układu odniesienia. Następnie poznaliśmy zasady dynamiki opisujące warunki, w jakich ciało pozostaje w spoczynku, jak też warunki powstawania różnych rodzajów ruchów. Nasuwa się pytanie, czy dokładnie te same prawa rządzą ruchami we wszystkich dowolnie wybranych układach odniesienia?

Postawmy zagadnienie inaczej. Załóżmy, że stwierdziliśmy stosowalność zasad dynamiki w odniesieniu do pewnego ruchu opisywanego względem układu związanego nieruchomo z Ziemią. Wywołajmy taki sam ruch np. w kabinie poruszającego się statku lub na obracającej się karuzeli. Czy prawa rządzące tym ruchem będą we wszystkich układach jednakowe?

Wprowadźmy następującą definicję. Taki *układ odniesienia, w którym stosują się zasady dynamiki Newtona, nazywamy układem inercjalnym.* Za taki układ będziemy uważali (na razie) Ziemię i wszelkie układy nieruchomo z nią związane.

Zbadamy kolejno następujące zagadnienia:

a) Czy układ poruszający się ruchem prostoliniowym jednostajnym względem układu inercjalnego jest też układem inercjalnym?

b) Czy układ poruszający się ruchem postępowym zmiennym względem układu inercjalnego jest układem inercjalnym?

c) Czy układ odbywający ruch obrotowy względem układu inercjalnego jest układem inercjalnym?

Rozważmy punkt a). Układ I utworzony np. przez przecięcie płaszczyzn ścian i podłogi pokoju jest z definicji układem inercjalnym. W pokoju przesuwany jest stół ruchem jednostajnym. Trzy krawędzie stołu stanowią układ II, który przesuwa się ruchem jednostajnym z prędkością $\mathbf{u}$ (prędkość unoszenia) względem układu I. W układzie II (np. po powierzchni stołu) przesuwa się badane ciało.

Niech w chwili t_1 prędkość tego ciała względem układu I wynosi $\mathbf{v}'_{\mathrm{I}}$, względem układu II — $\mathbf{v}'_{\mathrm{II}}$. Obowiązuje zależność

$$\mathbf{v}'_{\mathrm{I}} = \mathbf{v}'_{\mathrm{II}}+\mathbf{u},$$

gdyż $\mathbf{v}'_{\mathrm{I}}$ jest prędkością ruchu złożonego.

Podobnie w chwili t_2 spełnione jest równanie

$$\mathbf{v}''_{\mathrm{I}} = \mathbf{v}''_{\mathrm{II}}+\mathbf{u}.$$

Prędkości $\mathbf{v}''_{\mathrm{I}}$ i $\mathbf{v}''_{\mathrm{II}}$ nie muszą być odpowiednio równe prędkościom $\mathbf{v}'_{\mathrm{I}}$ i $\mathbf{v}'_{\mathrm{II}}$, gdyż nie robiliśmy założenia co do jednostajności ruchu ciała badanego. Jedynie prędkość $\mathbf{u}$ układu II jest niezmienna.

Odejmując stronami wypisane wyżej równania znajdujemy

$$\mathbf{v}''_{\mathrm{I}}-\mathbf{v}'_{\mathrm{I}} = \mathbf{v}''_{\mathrm{II}}-\mathbf{v}'_{\mathrm{II}}.$$

Dzieląc te przyrosty prędkości przez odpowiadający im czas t_2-t_1 i przechodząc do granicy znajdujemy przyspieszenie

$$\mathbf{a}_{\mathrm{I}} = \mathbf{a}_{\mathrm{II}}.$$

Jest to wynik bardzo ważny. Z równania tego wyciągamy następujące wnioski:

1. Jeśli badany ruch jest jednostajny względem układu inercjalnego I (czyli $\mathbf{a}_{I} = 0$), to jest też jednostajny względem układu II ($\mathbf{a}_{II} = 0$).

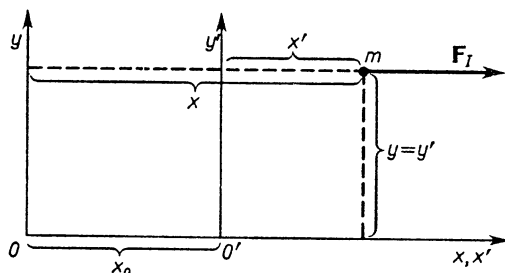
2. Jeśli ten ruch odbywa się z przyspieszeniem $\mathbf{a}_{I}$ względem układu inercjalnego I, to z takim samym przyspieszeniem odbywa się względem układu II.

W pierwszym przypadku zgodnie z zasadą bezwładności na ciało nie działa żadna siła albo wypadkowa wszystkich sił działających równa się zeru. W drugim przypadku działa siła równa iloczynowi masy i przyspieszenia, a więc jednakowa w układzie I i w układzie II:

$$\mathbf{F} = m\mathbf{a}_{I} = m\mathbf{a}_{II}.$$

Zbierając te wnioski powiemy, że w układzie II rządzą badanym ruchem te same prawa, co w układzie inercjalnym I. A zatem, *każdy układ poruszający się ruchem prostoliniowym jednostajnym względem układu inercjalnego jest również układem inercjalnym.*

Przechodzimy obecnie do punktu b). Czy układ II poruszający się ruchem zmiennym względem układu inercjalnego jest też układem inercjalnym?



Rys. 2.8

Matematycznie uprościmy zagadnienie biorąc pod uwagę dwa układy współrzędnych (rys. 2.8) o pokrywających się osiach Ox i $O'x'$ oraz równoległych osiach Oy i $O'y'$, jak również Oz i $O'z'$ (te ostatnie nie są zaznaczone na rysunku). Początki obu układów znajdują się w odległości x_0 od siebie. Należy pamiętać, że zgodnie z założeniem układ II odbywa ruch zmienny względem układu I, czyli x_0 nie jest stałe, lecz zmienia się z biegiem czasu

$$x_0 = f(t).$$

Niech ciało o masie m porusza się wzdłuż prostej równoległej do osi Ox pozostając w układzie inercjalnym xyz pod działaniem siły $\mathbf{F}_{I}$ skierowanej równolegle do osi Ox. Siła ta, zgodnie z II zasadą dynamiki, wywołuje w układzie I przyspieszenie

$$a_{I} = \frac{d^2x}{dt^2},$$

gdyż podczas ruchu ciała nie zmienia się ani współrzędna y, ani współrzędna z.

Spróbujemy inaczej wyrazić to przyspieszenie. Z rysunku 2.8 widać, że współrzęd-

na x, określająca położenie ciała m w dowolnej chwili, równa się sumie odległości x_0 w tejże chwili oraz współrzędnej x', określającej położenie ciała w II układzie:

$$x = x' + x_0.$$

Różniczkując względem czasu znajdujemy prędkość

$$\frac{dx}{dt} = \frac{dx'}{dt} + \frac{dx_0}{dt}.$$

Różniczkując po raz drugi znajdujemy przyspieszenie

$$\frac{d^2x}{dt^2} = \frac{d^2x'}{dt^2} + \frac{d^2x_0}{dt^2}.$$

W tym ostatnim wyrażeniu

$\frac{d^2x}{dt^2} = a_{\mathrm{I}}$ jest przyspieszeniem ciała w układzie inercjalnym I,

$\frac{d^2x'}{dt^2} = a_{\mathrm{II}}$ jest przyspieszeniem ciała w układzie II,

$\frac{d^2x_0}{dt^2} = a_u$ jest przyspieszeniem unoszenia, z jakim porusza się układ II względem układu inercjalnego.

Po podstawieniu tych oznaczeń otrzymujemy

$$a_{\mathrm{I}} = a_{\mathrm{II}} + a_u.$$

Jak widać, przyspieszenie ciała w układzie II jest inne niż przyspieszenie ciała w układzie inercjalnym. Z nierówności przyspieszeń wynika nierówność sił działających w układzie II i w układzie inercjalnym I:

$$ma_{\mathrm{II}} \neq ma_{\mathrm{I}},$$

a więc a_{II} nie może być uważane za wynik działania siły F_{I} zgodny z II zasadą dynamiki. Chcąc obliczyć przyspieszenie a_{II} należy dodatkowo uwzględnić przyspieszenie a_u, z jakim porusza się układ II:

$$a_{\mathrm{II}} = a_{\mathrm{I}} - a_u.$$

Mnożąc obie strony równania przez m znajdujemy

$$ma_{\mathrm{II}} = ma_{\mathrm{I}} + (-ma_u),$$

czyli

$$\mathbf{F}_{\mathrm{II}} = \mathbf{F}_{\mathrm{I}} + \mathbf{F}_0.$$

Treść tego równania możemy ująć w ten sposób: *Rozpatrując ruch ciała z punktu widzenia obserwatora znajdującego się w układzie nieinercjalnym, tzn. poruszającym się ruchem zmiennym względem układu inercjalnego, musimy do siły* $\mathbf{F}_{\mathrm{I}}$ *działającej na ciało w układzie inercjalnym dodawać siłę* $\mathbf{F}_0$ *równą liczbowo iloczynowi masy ciała przez przyspieszenie* $\mathbf{a}_u$ *układu, lecz skierowaną przeciwnie względem przyspieszenia układu:*

$$\mathbf{F}_0 = -m\mathbf{a}_u.$$

Siłę tę nazywamy *siłą oporu bezwładnego* albo krótko — *siłą bezwładności. Ciało spoczywa w układzie nieinercjalnym, gdy suma wszystkich sił działających, łącznie z siłą bezwładności, równa się zeru.* Jest to treść *zasady d'Alemberta.*

Występowanie siły oporu bezwładnego rozpatrzymy na kilku przykładach. Winda z pasażerem jedzie do góry ruchem jednostajnie przyspieszonym z przyspieszeniem $\mathbf{a}_u$. Wtedy na pasażera działa siła oporu bezwładnego $\mathbf{F}_0$ skierowana pionowo w dół. Ta siła dodaje się liczbowo do ciężaru pasażera $\mathbf{P}$, który tym samym silniej naciska na podłogę windy. W tym przypadku wartość liczbowa siły $\mathbf{F}$, z jaką pasażer działa na podłogę windy, wynosi

$$F = P+F_0.$$

Odwrotnie, jeśli winda jedzie w dół ruchem jednostajnie przyspieszonym z przyspieszeniem $\mathbf{a}_u$, to siła oporu bezwładnego $\mathbf{F}_0 = -m\mathbf{a}_u$ skierowana jest ku górze. Przy tym kierunku siły $\mathbf{F}_0$ suma geometryczna sił $\mathbf{P}$ i $\mathbf{F}_0$ sprowadza się do ich różnicy arytmetycznej i oddziaływanie pasażera na podłogę windy jest mniejsze od jego ciężaru:

$$F = P-F_0.$$

Rozważmy inny przykład. Na szynach znajduje się wózek w postaci idealnie gładkiej platformy. Na wózku znajduje się obserwator A nieruchomo związany z platformą. Na ziemi w pobliżu wózka znajduje się obserwator B. Na platformie spoczywa ciało o idealnie gładkiej podstawie. Wózek zostaje wprawiony w ruch z przyspieszeniem a_u. Ciało wobec braku tarcia nie utrzymuje się w tym samym miejscu powierzchni platformy, lecz pozostaje nadal w tym samym miejscu przestrzeni (w tym samym miejscu względem Ziemi, czyli w układzie inercjalnym). Platforma stopniowo wysuwa się spod ciała, które w końcu traci oparcie i spada na ziemię (lub też ciało dochodzi do tylnej ściany platformy i tam zatrzymuje się, bo suma siły reakcji ściany platformy i siły bezwładności równa się zeru). Ta ostatnia część zjawiska już nas nie interesuje, gdyż abstrahujemy od działania siły ciężkości. Takie będą spostrzeżenia obserwatora B. Spoczynek ciała względem układu inercjalnego, w którym się on znajduje, uzna on zgodnie z I zasadą dynamiki za wynik równoważenia się sił działających na ciało.

Opisane wyżej zachowanie się ciała zostanie zupełnie inaczej wyjaśnione przez obserwatora A poruszającego się wraz z platformą (może on sobie w ogóle nie zdawać sprawy z ruchu platformy). Jego zdaniem ciało zsuwa się z platformy ruchem jednostajnie przyspieszonym z pewnym przyspieszeniem. Po wyznaczeniu tego przyspieszenia okaże się, że jest ono równe $-\mathbf{a}_u$, czyli równe, lecz przeciwnie skierowane niż przyspieszenie ruchu platformy w układzie obserwatora B. Taki ruch ciała obserwator A przypisze działaniu siły $\mathbf{F}_0 = -m\mathbf{a}_u$, czyli właśnie działaniu siły bezwładności.

Nawiązując do równania

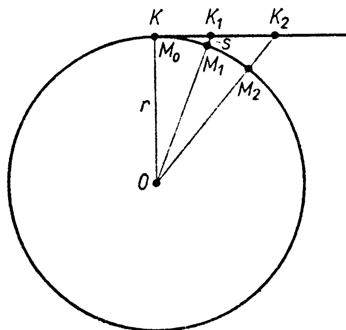
$$\mathbf{F}_{II} = \mathbf{F}_I+\mathbf{F}_0$$

podkreślamy raz jeszcze, że charakteryzując ruch ciała względem układu poruszającego się ruchem zmiennym w stosunku do układu inercjalnego musimy do siły działającej w układzie inercjalnym dodawać geometrycznie siłę bezwładności $\mathbf{F}_0$. Istnienia tej siły nie stwierdza obserwator nieruchomy z układu inercjalnego, w którym stosują się za-

sady dynamiki. A zatem nie można w sposób bezwzględny konstatować faktu istnienia lub nieistnienia siły, jest to bowiem zależne od układu odniesienia.

Z tych rozważań wynika następujący wniosek ogólny. *Układ poruszający się ruchem zmiennym względem układu inercjalnego nie jest układem inercjalnym.* Różni się on od układu inercjalnego tym, że występują w nim dodatkowe siły — siły bezwładności.

Zajmiemy się obecnie punktem c). Przechodzimy do badania zachowania się ciał w układzie obracającym się. Rozważymy oddzielnie przypadek, gdy ciało spoczywa w układzie obracającym się i przypadek, gdy ciało porusza się w układzie obracającym się. Niech naszym układem obracającym się będzie np. pozioma tarcza wirująca jednostajnie dokoła osi pionowej. Nieruchomo z tarczą związany jest obserwator znajdujący się w pobliżu jej obwodu. Obok obserwatora spoczywa kula. Tarcza zostaje wprawiona w ruch obrotowy. W pewnej chwili zabrakło siły dośrodkowej utrzymującej kulę na tarczy, wobec czego kula zsuwa się z tarczy. W ciągu krótkiego czasu, który potem następuje, ruch kuli inaczej opisuje obserwator A odbywający w dalszym ciągu ruch wraz z tarczą, a inaczej obserwator B stojący poza tarczą na ziemi. Obserwator B mówi, że kula posuwa się po stycznej $M_0K_1K_2$ (rys. 2.9) zachowując tę prędkość liniową $\mathbf{v}$,



Rys. 2.9

którą miała będąc na tarczy w punkcie M_0. Obserwator A porusza się wraz z tarczą zajmując kolejno w przestrzeni po równych, bardzo krótkich odstępach czasu położenia M_0, M_1, M_2 itd. Stale mu się wydaje, że kula oddala się od niego wzdłuż przedłużeń promieni tarczy OM_1, OM_2, ... Po bardzo krótkim czasie τ niech obserwator A będzie w M_1, a kula w punkcie K_1. Szukamy drogi $s = K_1M_1$ przebytej przez kulę w stosunku do obserwatora w czasie τ.

Można wypisać zależność

$$K_1O^2 = K_1M_0^2 + OM_0^2,$$

ale $K_1O = K_1M_1 + OM_1$, czyli $K_1O = s+r$; $K_1M_0 = v\tau$ jako droga kuli przebyta ruchem jednostajnym z prędkością v wzdłuż stycznej. Po podstawieniu tych wartości otrzymujemy

$$(s+r)^2 = v^2\tau^2 + r^2,$$

$$s^2 + 2sr + r^2 = v^2\tau^2 + r^2.$$

Droga s przebyta w bardzo krótkim czasie τ jest bardzo mała. Wielkość s^2 można więc zaniedbać. A zatem

$$s = \frac{1}{2}\frac{v^2}{r}\tau^2.$$

Otrzymaliśmy wyrażenie na drogę typu

$$s = \frac{1}{2}at^2,$$

z tym że a zastąpione jest przez v^2/r. Stąd wnioskujemy, że ciało odsuwa się od obserwatora pozostającego na wirującej tarczy ruchem jednostajnie zmiennym z przyspieszeniem

$$a = \frac{v^2}{r},$$

a więc pod działaniem siły

$$F = m\frac{v^2}{r}.$$

Wobec założenia, że czas jest bardzo krótki, możemy wnioskować, że kula już wtedy, gdy spoczywała w układzie wirującym, podlegała działaniu siły skierowanej wzdłuż promienia tarczy od środka na zewnątrz. W badanej sile poznajemy wprowadzoną w poprzednim paragrafie siłę odśrodkową bezwładności. Przekonaliśmy się, że *siła odśrodkowa bezwładności równa się liczbowo sile dośrodkowej* $\frac{mv^2}{r}$ *lub* $m\omega^2 r$, *gdzie* ω *oznacza prędkość kątową układu wirującego.*

Tak więc możemy powiedzieć, że z punktu widzenia obserwatora ruchomego ciało spoczywające w układzie obracającym się z prędkością kątową ω podlega dodatkowej (nie istniejącej w układzie inercjalnym) sile odśrodkowej bezwładności.

Obserwator A trzymający kulę w ręku dostarcza siły dośrodkowej $\mathbf{F}_I$ potrzebnej do utrzymania kuli w ruchu obrotowym wraz z tarczą względem układu inercjalnego. Do charakterystyki zachowania się kuli względem układu wirującego ta jedna siła nie wystarcza. Prócz niej musi być uwzględniona siła odśrodkowa bezwładności $\mathbf{F}_0$. Obie siły jako sobie równe i przeciwnie skierowane, działające na to samo ciało, równoważą się wzajemnie i dlatego kula spoczywa względem układu wirującego (zasada d'Alemberta).

Wiemy, że Ziemia odbywa też ruch obrotowy dokoła własnej osi. Obliczmy, jak wielka jest siła odśrodkowa bezwładności działająca na ciało o masie m znajdujące się na równiku i odbywające wraz z Ziemią ruch obrotowy. W zaokrągleniu przyjmujemy, że promień Ziemi $R_z = 6{,}3 \cdot 10^6$ m:

$$\omega = \frac{2\pi}{T} = \frac{6{,}28 \text{ rad}}{86\,400 \text{ s}} \simeq 7{,}3 \cdot 10^{-5} \text{ rad} \cdot \text{s}^{-1}.$$

Stąd

$$F_{\text{odśr bezwł}} \simeq 0{,}0336\,\{m\}\ \text{N},$$

gdzie symbol $\{m\}$ oznacza wartość liczbową masy ciała w układzie SI (p. U.1.1).

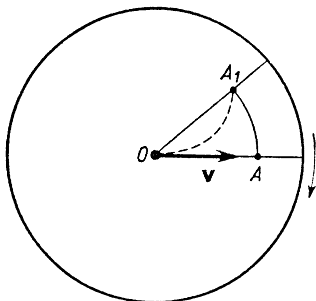
Porównajmy tę siłę z ciężarem tegoż ciała P:

$$P = mg = 9{,}78\,\{m\}\,\mathrm{N},$$

gdzie symbol $\{m\}$ ma to samo znaczenie, co poprzednio.

Siła odśrodkowa bezwładności stanowi na równiku ziemskim zaledwie około $\frac{1}{3}\%$ ciężaru ciała. Zaniedbujemy tę siłę traktując Ziemię jako układ inercjalny.

Obecnie rozważymy ruch punktu materialnego w odniesieniu do układu wirującego, np. w stosunku do tarczy poziomej wirującej dokoła osi pionowej (rys. 2.10). Znów uprościmy zagadnienie zakładając, że punkt materialny znajdujący się początkowo nad środkiem tarczy przesuwa się nad nią w kierunku równoległym do jednego z promieni ruchem jednostajnym z prędkością $\mathbf{v}$. Tarcza wiruje ze stałą prędkością kątową $\boldsymbol{\omega}$. Jak mówiliśmy już poprzednio, prędkość kątową traktujemy jako wektor skierowany wzdłuż osi obrotu zgodnie z regułą śruby prawoskrętnej. Jeżeli, jak na rys. 2.10, kierunek obrotu tarczy jest zgodny z ruchem wskazówek zegara, prędkość kątowa $\boldsymbol{\omega}$ jest skierowana za płaszczyznę rysunku. *Prędkość* $\mathbf{v}$ badanego punktu materialnego, jak wynika z poprzedniego założenia, jest prostopadła do wektora $\boldsymbol{\omega}$.



Rys. 2.10

Niech punkt materialny posuwa się przez krótki czas τ. W tym czasie odsunie się on od środka tarczy o odległość $r = v\tau$. Gdyby tarcza była nieruchoma, to po tym czasie znalazłby się on nad punktem A. Ponieważ jednak tarcza wiruje, więc na miejsce punktu A po czasie τ przesunie się punkt A_1, odległy od A o łuk AA_1:

$$AA_1 = \omega\tau r,$$

$$AA_1 = \omega\tau v\tau,$$

$$AA_1 = \omega v\tau^2.$$

Jakie będą spostrzeżenia obserwatora znajdującego się w punkcie A na tarczy?

Dla uzyskania większej poglądowości wyobraźmy sobie, że nasz punkt materialny jest kulą wylatującą z rewolweru umieszczonego nad środkiem tarczy. Obserwator widzi wyraźnie, że w chwili strzału lufa rewolwerowa skierowana jest wprost ku niemu. Po czasie τ, potrzebnym na przebycie przez kulę odcinka równego OA, kula nie trafia w obserwatora, lecz przelatuje bokiem nad punktem A_1. Punkt A_1 jest przesunięty względem punktu A o łuk AA_1 w kierunku przeciwnym względem ruchu obrotowego

tarczy. Innymi słowy, obserwator stwierdza, że kula wykonała w czasie τ ruch złożony. Jeden z ruchów składowych jest jednostajny i odbywa się z prędkością v wzdłuż promienia, drugi odbywa się po łuku AA_1 przy czym $AA_1 = v\omega\tau^2$. Przekształcamy to wyrażenie tak, aby uzyskać analogię do wzoru $s = \frac{1}{2}at^2$:

$$\breve{AA}_1 = \frac{1}{2}2v\omega\tau^2.$$

Drugi ruch składowy jest więc jednostajnie zmienny i odbywa się z przyspieszeniem $a = 2v\omega$. Obserwator na tarczy przypisuje ten ruch działaniu dodatkowej siły

$$F_C = m \cdot 2v\omega,$$

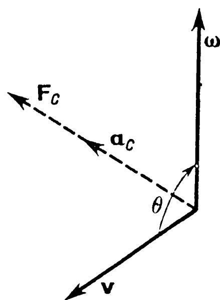
zwanej *siłą Coriolisa.* Siła ta ma kierunek prostopadły zarówno do prędkości $\mathbf{v}$, jak i do $\boldsymbol{\omega}$. Można udowodnić, że zawsze, przy dowolnym kącie między wektorami $\mathbf{v}$ i $\boldsymbol{\omega}$, przyspieszenie Coriolisa

$$\mathbf{a}_C = 2\mathbf{v}\times\boldsymbol{\omega}, \tag{2.6}$$

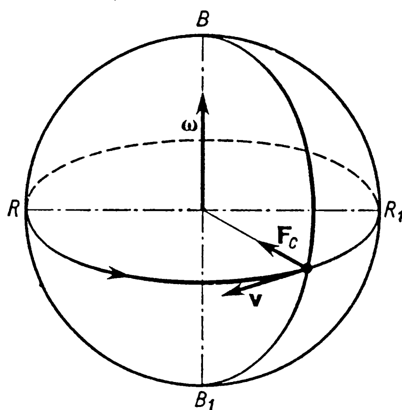
a siła Coriolisa

$$\mathbf{F}_C = 2m\mathbf{v}\times\boldsymbol{\omega}. \tag{2.7}$$

Rysunek 2.11 przedstawia zestawienie wektorów $\boldsymbol{\omega}$, $\mathbf{v}$, $\mathbf{a}_C$ i $\mathbf{F}_C$. Wektory $\mathbf{a}_C$ i $\mathbf{F}_C$ są prostopadłe do płaszczyzny zawierającej wektory $\mathbf{v}$ i $\boldsymbol{\omega}$, tworzące w ogólnym przypadku dowolny kąt θ.



Rys. 2.11



Rys 2.12

Rozważmy siłę Coriolisa działającą na ciało poruszające się względem Ziemi. Wiemy, że Ziemia odbywa ruch obrotowy jednostajny dokoła własnej osi w kierunku z zachodu na wschód z prędkością kątową $\omega = 7{,}3 \cdot 10^{-5}$ rad $\cdot$ s^{-1}. Jaka siła Coriolisa będzie działała na samolot o masie m kilogramów, lecący np. nad równikiem w kierunku ze wschodu za zachód (rys. 2.12)? Prędkość samolotu niech będzie równa 360 km/h = 100 m/s. Wobec prostopadłości kierunków $\mathbf{v}$ i $\boldsymbol{\omega}$

$$F_c = 2mv\omega,$$

skąd po podstawieniu danych liczbowych w układzie SI otrzymujemy przybliżoną wartość

$$F_c = 0{,}0146\{m\}\ \mathrm{N},$$

gdzie symbol $\{m\}$ oznacza wartość liczbową masy samolotu w układzie SI. Stanowi to około 1,5 promille ciężaru samolotu. Nic dziwnego, że wobec tak niewielkiej stosunkowo wartości wpływ siły Coriolisa na ruch ciał na Ziemi został niezbyt dawno stwierdzony.

Siła Coriolisa działająca na omawiany wyżej samolot, jak wynika z wiadomości o iloczynie wektorowym, skierowana jest do środka Ziemi, tzn. przyczynia się do pozornego zwiększenia ciężaru samolotu. Gdyby samolot leciał z zachodu na wschód, zwrot siły Coriolisa byłby odwrotny, tzn. pionowy w górę.

Siła Coriolisa zaznacza się wyraźniej przy ruchu pocisków w związku z ich dużymi prędkościami: prędkość kuli karabinowej około 800 m/s, prędkość pocisków artyleryjskich około 1700 m/s. Porównując te prędkości z prędkością samolotu w naszym przykładzie, stwierdzimy w przypadku kuli karabinowej 8-krotne, a w przypadku pocisku artyleryjskiego 17-krotne powiększenie przyspieszenia Coriolisa, gdy kierunek prędkości pocisku jest z zachodu na wschód lub odwrotnie.

Siła Coriolisa ma wpływ na kształtowanie się brzegów rzek. Weźmy pod uwagę rzekę płynącą na półkuli północnej wzdłuż południka. Łatwo stwierdzić, opierając się na właściwościach iloczynu wektorowego, że pod działaniem siły Coriolisa na półkuli północnej zawsze jest podmywany brzeg prawy, a na półkuli południowej — brzeg lewy. Te same wyniki otrzymamy rozważając bieg rzek płynących w dowolnych kierunkach na obu półkulach, jedynie wartość siły Coriolisa będzie zmienna w zależności od kąta między prędkością rzeki $\mathbf{v}$ i prędkością kątową Ziemi $\boldsymbol{\omega}$. Tylko na samym równiku rzeki przecinające równik pod kątem prostym nie podlegają sile Coriolisa wobec równoległości wektorów $\mathbf{v}$ rzeki i $\boldsymbol{\omega}$ Ziemi.

Na podstawie tych rozważań możemy powiedzieć, że nasze początkowe założenie, że układy związane nieruchomo z Ziemią są układami inercjalnymi, jest spełnione tylko w przybliżeniu. Ziemia sama odbywa ruch obrotowy, ale, jak stwierdziliśmy, występująca na Ziemi siła odśrodkowa bezwładności i siła Coriolisa są niewielkie.

Układem inercjalnym jest układ związany z gwiazdami stałymi. W takim układzie dotychczas nie odkryto odstępstw od praw dynamiki. Wszystkie układy odbywające ruch jednostajny prostoliniowy względem układu inercjalnego są też układami inercjalnymi.

Czy można twierdzić, że pewien układ inercjalny jest w spoczynku, a inny w ruchu? Badanie ruchu ciał w tych układach nie daje nam do tego żadnych podstaw. Nie możemy powiedzieć, że którykolwiek z układów jest w bezwzględnym spoczynku, a inny w ruchu.

Ani bezwzględny spoczynek, ani bezwzględny ruch nie może być wykryty — na tym polega *zasada względności ruchów mechaniki klasycznej*.

Pytania. 1. Czy szerokość geograficzna określająca położenie danego punktu na powierzchni Ziemi ma wpływ na wartość siły odśrodkowej bezwładności? (rozumowanie poprzeć rysunkiem).

2. Przy jakim kierunku prędkości $\mathbf{v}$, niezależnie od jej wartości liczbowej, siła Coriolisa w układzie obracającym się z prędkością kątową $\boldsymbol{\omega}$ jest równa zeru? Jakim warunkom ruchu na Ziemi to odpowiada?

2.7. Dynamika układu punktów materialnych. Środek masy

W dotychczasowych naszych rozważaniach dotyczących dynamiki ruchu postępowego traktowaliśmy ciała poruszające się jako punkty materialne. Obecnie przejdziemy do dynamiki układu punktów materialnych, przy czym warto od razu podkreślić, że ciało modelowe o ciągłym rozkładzie masy również może być traktowane jako układ nieskończonej liczby punktów materialnych.

Rozpoczniemy od zdefiniowania pojęcia tzw. środka masy. *Środek masy układu n punktów materialnych* o masach $m_1, m_2, \ldots, m_n$ jest punktem, którego współrzędne x_0, y_0, z_0 w danym układzie współrzędnych wyrażają się następującymi wzorami:

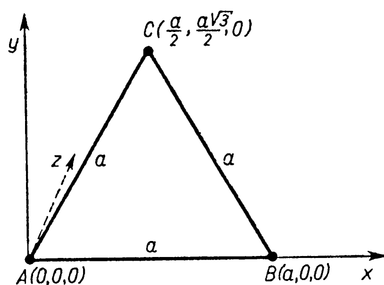
$$x_0 = \frac{\sum m_i x_i}{\sum m_i}, \qquad y_0 = \frac{\sum m_i y_i}{\sum m_i}, \qquad z_0 = \frac{\sum m_i z_i}{\sum m_i}, \tag{2.8}$$

gdzie i zmienia się od 1 do n.

Zastosujmy te definicyjne równania do wyznaczenia współrzędnych środka masy dwóch punktów materialnych m_1 i m_2. Dobierzmy tak układ współrzędnych, aby oba punkty leżały na osi x i miały współrzędne $(x_1, 0, 0)$ i $(x_2, 0, 0)$. W tych warunkach

$$x_0 = \frac{m_1 x_1 + m_2 x_2}{m_1 + m_2}, \qquad y_0 = 0, \qquad z_0 = 0.$$

W przypadku układu trzech punktów materialnych rozmieszczonych w określony sposób, np. w wierzchołkach równobocznego trójkąta, postępujemy następująco: 1) dobieramy układ współrzędnych możliwie najdogodniejszy do dalszych obliczeń, np. taki jak na rys. 2.13, 2) ustalamy współrzędne wierzchołków trójkąta, 3) wypisujemy równania na x_0, y_0, z_0.



Rys. 2.13

Podczas obliczania współrzędnych środka masy ciała sztywnego zakładamy jego budowę ciągłą i dzielimy je w myśli na nieskończenie wiele ($n \to \infty$) małych elementów o masach Δm_i. Wtedy wzory na współrzędne środka masy przyjmują postać

$$x_0 = \lim_{\Delta m_i \to 0} \frac{\sum \Delta m_i x_i}{\sum \Delta m_i}, \qquad y_0 = \lim_{\Delta m_i \to 0} \frac{\sum \Delta m_i y_i}{\sum \Delta m_i}, \qquad z_0 = \lim_{\Delta m_i \to 0} \frac{\sum \Delta m_i z_i}{\sum \Delta m_i}.$$

Z matematyki wiadomo, że granice takich sum można zastąpić całkami. A zatem w odniesieniu do ciał o budowie ciągłej obowiązują równania:

$$x_0 = \frac{\int x\,dm}{\int dm} = \frac{\int x\,dm}{m_0},$$
$$y_0 = \frac{\int y\,dm}{\int dm} = \frac{\int y\,dm}{m_0}, \qquad (2.9)$$
$$z_0 = \frac{\int z\,dm}{\int dm} = \frac{\int z\,dm}{m_0},$$

gdzie $m_0 = \int dm$ jest całkowitą masą ciała.

Istnienie elementów symetrii, jak środek, oś lub płaszczyzna symetrii, jednorodnych brył upraszcza ustalanie położenia środka masy, gdyż leży on zawsze na elemencie symetrii. Tak np. środek masy jednorodnej kuli leży w jej środku, środek masy jednorodnego walca leży na jego osi symetrii itd.

Z dotychczasowych rozważań — o charakterze raczej matematycznym — nie widać jeszcze korzyści, jakie wynikają z wprowadzenia pojęcia środka masy. Korzyści te wyraźnie się zaznaczą, gdy zajmiemy się ruchem postępowym ciała złożonego z szeregu punktów materialnych o łącznej masie m_0. Współrzędna x_0 środka masy spełnia równanie

$$m_0 x_0 = m_1 x_1 + m_2 x_2 + m_3 x_3 + \ldots,$$

przy czym, wobec tego że ciało jest w ruchu, każda ze współrzędnych $x_0, x_1, x_2, x_3, \ldots$ jest funkcją czasu. Różniczkując względem czasu znajdujemy zależność między prędkościami

$$m_0 \frac{dx_0}{dt} = m_1 \frac{dx_1}{dt} + m_2 \frac{dx_2}{dt} + m_3 \frac{dx_3}{dt} + \ldots,$$

$$m_0 v_{0x} = m_1 v_{1x} + m_2 v_{2x} + m_3 v_{3x} + \ldots, \qquad (2.10)$$

gdzie v_{0x} oznacza składową prędkości środka masy ciała w kierunku osi x. Różniczkując drugi raz znajdujemy związek między przyspieszeniami:

$$m_0 \frac{dv_{0x}}{dt} = m_1 \frac{dv_{1x}}{dt} + m_2 \frac{dv_{2x}}{dt} + m_3 \frac{dv_{3x}}{dt} + \ldots,$$

$$m_0 a_{0x} = m_1 a_{1x} + m_2 a_{2x} + m_3 a_{3x} + \ldots,$$

gdzie a_{0x} oznacza składową przyspieszenia środka masy ciała w kierunku osi x.

Uwzględniając drugą zasadę dynamiki można powyższe równanie przepisać w postaci

$$m_0 a_{0x} = F_{1x} + F_{2x} + F_{3x} + \ldots,$$

gdzie $F_{1x}, F_{2x}, F_{3x}, \ldots$ są składowymi x-owymi sił wypadkowych, działających odpowiednio na masy m_1, m_2, m_3 itd.

Dla obu pozostałych osi można wypisać równania analogiczne. Od tych trzech równań skalarnych można przejść do jednego równania wektorowego:

$$m_0 \mathbf{a}_0 = \mathbf{F}_1 + \mathbf{F}_2 + \mathbf{F}_3 + \ldots = \sum \mathbf{F}.$$

A zatem *środek masy ciała ma tę właściwość, że iloczyn całkowitej masy* m_0 *i przyspieszenia środka masy* $\mathbf{a}_0$ *równa się sumie geometrycznej wszystkich sił działających na poszczególne punkty układu. Siły* te można by podzielić na *zewnętrzne* $\mathbf{F}_z$ (tzn. działające między punktami układu i punktami znajdującymi się zewnątrz rozważanego układu) i *wewnętrzne* $\mathbf{F}_w$ (tzn. działające między punktami danego układu):

$$\mathbf{F} = \sum \mathbf{F}_z + \sum \mathbf{F}_w. \tag{2.11}$$

Z trzeciej zasady dynamiki wynika, że siły wewnętrzne występują parami, których składniki są równe co do wartości, lecz przeciwne co do kierunku. Stąd wniosek, że wypadkowa wszystkich sił wewnętrznych równa się zeru i wspomnianą właściwość środka masy można wyrazić prostszą zależnością

$$m_0 \mathbf{a}_0 = \sum \mathbf{F}_z. \tag{2.12}$$

Innymi słowy, *środek masy ciała porusza się tak, jakby w nim była skupiona całkowita masa poddana działaniu wypadkowej wszystkich sił zewnętrznych.* Stwierdzenie powyższe jest słuszne zarówno w odniesieniu do układu sztywnego o niezmiennych wzajemnych odległościach poszczególnych cząstek, jak również dla układu, w którego skład wchodzą cząstki wykonujące dowolne ruchy pod wpływem sił wewnętrznych.

2.8. Zasada zachowania pędu

W odniesieniu do pojedynczego punktu materialnego drugą zasadę dynamiki zapisaliśmy w postaci (2.3)

$$\mathbf{F} = \frac{d\mathbf{p}}{dt}.$$

Nasuwa się pytanie, jaki jest odpowiednik tego podstawowego równania dynamiki dla układu punktów materialnych.

Wracając do równania (2.10) stwierdzimy, że iloczyn całkowitej masy układu przez składową prędkości środka masy w kierunku osi x, czyli składowa x-owa pędu środka masy p_{0x} równa się sumie składowych x-owych pędów poszczególnych mas układu

$$p_{0x} = p_{1x} + p_{2x} + p_{3x} + \dots$$

Uwzględniając analogiczne równania dla pozostałych osi współrzędnych można od trzech równań skalarnych przejść do jednego wektorowego:

$$\mathbf{p}_0 = \mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3 + \dots = \mathbf{p}_w, \tag{2.13}$$

gdzie $\mathbf{p}_0$ oznacza pęd środka masy układu równy $m_0\mathbf{v}_0$, $\mathbf{p}_w$ — pęd wypadkowy układu.

Treść równania (2.13) można ująć następująco: *pęd środka masy układu (czyli iloczyn całkowitej masy układu i prędkości środka masy) równa się pędowi wypadkowemu układu (czyli sumie geometrycznej pędów poszczególnych jego punktów materialnych).*

Zróżniczkowanie (2.13) względem czasu prowadzi do zależności

$$\frac{d\mathbf{p}_0}{dt} = \frac{d\mathbf{p}_1}{dt} + \frac{d\mathbf{p}_2}{dt} + \frac{d\mathbf{p}_3}{dt} + \dots \tag{2.14}$$

lub

$$\frac{d\mathbf{p}_0}{dt} = \frac{d\mathbf{p}_w}{dt}. \tag{2.15}$$

Uwzględniając (2.3) stwierdzimy, że każdy ze składników sumy po prawej stronie równania (2.14) przedstawia siłę wypadkową $\mathbf{F}_1$, $\mathbf{F}_2$, $\mathbf{F}_3$, ... działającą na punkty materialne $m_1, m_2, m_3, \ldots$ układu. A zatem

$$\sum \mathbf{F} = \frac{d\mathbf{p}_0}{dt} = \frac{d\mathbf{p}_w}{dt}.$$

Ponieważ jednak na podstawie trzeciej zasady dynamiki $\sum \mathbf{F}_w = 0$, więc ostatnie równanie redukuje się do postaci

$$\sum \mathbf{F}_z = \frac{d\mathbf{p}_0}{dt} = \frac{d\mathbf{p}_w}{dt}. \tag{2.16}$$

Zależności te są odpowiednikiem równania (2.3) obowiązującym w przypadku ruchu postępowego układu punktów materialnych. Treść równania (2.16) można wyrazić następująco: *wypadkowa wszystkich sił zewnętrznych działających na układ punktów materialnych równa się pochodnej względem czasu pędu środka masy lub pochodnej względem czasu wypadkowego pędu układu.*

Z równania (2.16) wynika bardzo ważna *zasada zachowania pędu*. Załóżmy, że wypadkowa wszystkich sił zewnętrznych działających na rozważany układ równa się zeru. Wtedy

$$\frac{d\mathbf{p}_w}{dt} = 0, \qquad \text{czyli} \qquad \mathbf{p}_w = \text{const.}$$

Innymi słowy, *gdy wypadkowa wszystkich sił zewnętrznych działających na układ równa się zeru, to wektor wypadkowego pędu całego układu pozostaje stały.* Zmiana pędu układu może być wywołana jedynie działaniem takich sił zewnętrznych, które się nawzajem nie równoważą. Żadne siły wewnętrzne nie są w stanie zmienić wypadkowego pędu układu.

Wypowiedzianej wyżej zasady nie należy rozumieć w ten sposób, że poszczególne punkty materialne wchodzące w skład układu nie mogą zmieniać swego pędu pod działaniem sił wewnętrznych. Wszak pęd wypadkowy

$$\mathbf{p}_w = \mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3 + \ldots$$

Poszczególne pędy $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$, ... pod działaniem sił wewnętrznych mogą ulegać zmianie, ale zawsze w ten sposób, by pęd wypadkowy układu pozostawał stały co do wartości liczbowej i kierunku.

Zasadę zachowania pędu zilustrujemy kilkoma przykładami. Wyskakując z łódki stojącej przy brzegu jeziora uzyskujemy pęd skierowany w stronę lądu. Równocześnie łódka — zgodnie z zasadą zachowania pędu — oddala się nieco od brzegu uzyskując pęd równy co do wartości, lecz przeciwnie skierowany. Wypadkowy pęd układu łódka-człowiek pozostaje nadal równy zeru. (Czy prędkość łódki jest równa prędkości wyskakującego człowieka?)

Pocisk wyrzucony z działa ukośnie względem poziomu odbywa ruch po torze (w przybliżeniu) parabolicznym. W pewnej chwili pod wpływem sił wewnętrznych pocisk ulega rozerwaniu. Ponieważ siły wewnętrzne nie zmieniają wypadkowego pędu układu, więc odłamki rozlatują się na wszystkie strony w ten sposób, że suma geometryczna ich pędów w chwili rozerwania jest równa pędowi pocisku tworzącego jeszcze całość. Za chwilę jednak sytuacja ulega zmianie, gdyż zarówno pocisk przed rozerwaniem, jak i odłamki z niego powstałe pozostają pod działaniem zewnętrznej siły przyciągania ziemskiego. Układ odłamków porusza się nadal tak, jakby cała jego masa była zebrana w środku masy i była poddana działaniu wypadkowej siły zewnętrznej, tzn. środek masy układu odłamków odbywa w dalszym ciągu ruch po pierwotnym torze parabolicznym.

Na zasadzie zachowania pędu opiera się działanie śruby okrętowej i śmigła samolotu. Śruba odrzuca wodę do tyłu, statek uzyskuje pęd skierowany ku przodowi. Podobnie śmigło odrzuca do tyłu masy powietrza, a samolot przesuwa się naprzód.

Znane są ogólnie zjawiska „odrzutu" przy użyciu broni palnej: dubeltówka czy karabin „uderzają" strzelca, lufa cofa się przy wystrzale. Zjawisko odrzutu jest wykorzystywane na szeroką skalę w samolotach odrzutowych i pociskach rakietowych. Zasada ich ruchu polega na tym, że w specjalnej komorze wewnętrznej odbywa się spalanie mieszanki wybuchowej. Gazy z dużą prędkością, a więc i z dużym pędem, uchodzą przez otwór w tylnej części samolotu lub rakiety, które równocześnie uzyskują pęd równy co do wartości, lecz skierowany ku przodowi.

Pytania i zadania

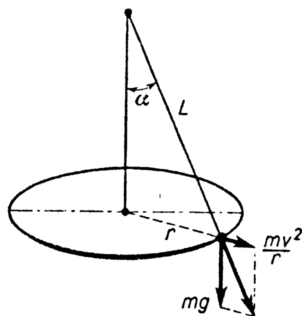
1. Czym się zajmuje dynamika? Jak można sformułować podstawowe zagadnienia dynamiki?
2. Co to jest bezwładność ciała? Co jest miarą bezwładności w ruchu postępowym?
3. Jakimi równaniami można wyrazić drugą zasadę dynamiki?
4. Podać definicję jednostki siły w układzie SI. Omówić sposoby pomiaru siły.
5. Omówić przykłady ilustrujące trzecią zasadę dynamiki.
6. Jakie siły występują w ruchu jednostajnym i niejednostajnym po okręgu?
7. Które z dwóch niżej podanych wyrażeń wektorowych określa siłę dośrodkową:

$$\mathbf{F} = -m\omega^2\mathbf{r} \quad \text{czy} \quad \mathbf{F} = m\omega^2\mathbf{r}?$$

8. Jak można zdefiniować współrzędne środka masy układu punktów materialnych?
9. Jakie właściwości ma środek masy układu z punktu widzenia dynamiki ruchu postępowego tego układu?
10. W jaki sposób dochodzi się do zasady zachowania pędu i jak można ściśle ująć jej treść?
11. Zilustrować przykładami zasadę zachowania pędu.
12. Jaka jest podstawowa różnica w opisie danego ruchu podanym przez obserwatora znajdującego się a) w inercjalnym układzie odniesienia, b) w nieinercjalnym układzie odniesienia?
13. Jakie dodatkowe siły należy uwzględniać w obracającym się układzie? Omówić dokładnie te siły w odniesieniu do ciał a) spoczywających na powierzchni Ziemi, b) poruszających się względem Ziemi.
14. Dlaczego w wielu przypadkach na Ziemi zaniedbuje się siłę odśrodkową bezwładności i siłę Coriolisa?
15. Wykazać, że w przypadku ruchu (zwanego harmonicznym), opisanego zależnością $x =$

$= A\cos(\omega t+\varphi)$, gdzie A, ω i φ są wielkościami stałymi, siła działająca jest proporcjonalna do wychylenia od położenia równowagi ($x = 0$) i skierowana przeciwnie do wychylenia.

16. Pasażer znajduje się w pociągu o zasłoniętych oknach, jadącym po prostoliniowym torze. Na podstawie jakich doświadczeń mógłby on wnioskować o ruchu a) jednostajnym, b) jednostajnie zmiennym, c) niejednostajnie zmiennym pociągu?



Rys. 2.14

17. Kulka wahadła stożkowego (rys. 2.14) o długości L wykonuje ruch jednostajny po okręgu w płaszczyźnie poziomej z prędkością v. Wyznaczyć kąt odchylenia nici od pionu.

$$\text{Odp. } \cos\alpha = \frac{-v^2+\sqrt{v^4+4g^2L^2}}{2gL}.$$

18. Małe ciało umieszczone w najwyższym punkcie powierzchni doskonale gładkiej kuli o promieniu R zaczyna się zsuwać po jej powierzchni. Na jakiej wysokości h liczonej względem poziomu początkowego punktu ruchu ciało oderwie się od powierzchni kuli?

Odp. $R/3$.

19. W miejscowości o szerokości geograficznej 45° spada swobodnie na Ziemię z wysokości $h = 500$ m ciało o masie $m = 10$ kg. Obliczyć siłę odśrodkową bezwładności i siłę Coriolisa, gdy ciało znajdzie się tuż nad powierzchnią Ziemi. Promień Ziemi przyjąć równy 6370 km.

Odp. 0,238 N, 0,104 N.

20. Wykazać, że środek masy dwóch punktów materialnych o masach m_1 i m_2 leży na prostej łączącej te punkty, przy czym jego odległości r_1 i r_2 od obu punktów spełniają proporcję

$$m_1 : m_2 = r_2 : r_1.$$

ROZDZIAŁ 3

Praca i energia

3.1. Praca

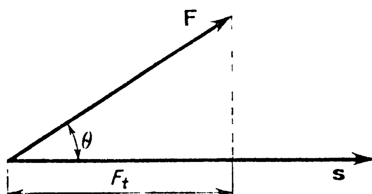
Rozważania dotyczące pracy w ścisłym fizycznym znaczeniu rozpoczniemy od przypadku, gdy pod działaniem stałej siły **F** punkt materialny odbywa prostoliniowe przesunięcie **s** (wektor przesunięcia **s** jest równy liczbowo przebytej drodze i ma kierunek zgodny z kierunkiem ruchu). *Pracę* W definiujemy jako iloczyn skalarny wektora siły **F** i wektora przesunięcia **s**, a zatem

$$W = \mathbf{F} \cdot \mathbf{s} \tag{3.1}$$

lub

$$W = Fs\cos\theta, \tag{3.2}$$

gdzie θ oznacza kąt między kierunkiem siły i przesunięcia (rys. 3.1). Zgodnie z podaną definicją *praca jest skalarem.*



Rys. 3.1

Jednostką pracy w układzie SI jest dżul (J)*. *Dżul* jest to praca wykonana podczas przesunięcia punktu materialnego pod działaniem siły 1 niutona na odległość 1 metra, gdy kierunki siły i przesunięcia są zgodne:

$$1\ \mathrm{J} = 1\ \mathrm{N} \cdot 1\ \mathrm{m} = 1\ \mathrm{kg} \cdot \mathrm{m^2/s^2}.$$

---

* Dżul jest równocześnie jednostką pracy w układzie MKSA. W układzie CGS jednostką pracy jest *erg*.

$$1\ \mathrm{erg} = 1\ \mathrm{dyna} \cdot 1\ \mathrm{cm} = 10^{-7}\ \mathrm{J}.$$

Jednostką pracy w układzie ciężarowym jest kilogramometr (kGm).

$$1\ \mathrm{kGm} = 9{,}80665\ \mathrm{J}.$$

Wróćmy do wzoru (3.2). Wynika z niego bezpośrednio, że praca może przyjmować zarówno wartości dodatnie, jak i ujemne. Praca danej siły jest dodatnia, gdy kąt θ jest ostry, ujemna — gdy kąt θ jest rozwarty i równa zeru nie tylko wtedy, gdy siła lub przesunięcie są równe zeru, lecz także, gdy kierunek siły jest prostopadły do kierunku przesunięcia. Tak np. praca wykonana przez siłę ciężkości jest dodatnia przy spadku ciała, ujemna — przy podnoszeniu go do góry i równa zeru — przy przesuwaniu ciała po torze poziomym.

Z rysunku (3.1) widać, że iloczyn $F\cos\theta$ przedstawia rzut siły $\mathbf{F}$ na kierunek przesunięcia $\mathbf{s}$, czyli jest wartością liczbową siły F_t stycznej do toru:

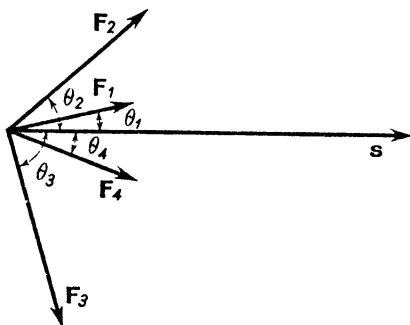
$$F\cos\theta = F_t,$$

a zatem

$$W = F_t s.$$

Jeżeli ciało porusza się pod działaniem kilku równocześnie działających sił $\mathbf{F}_1$, $\mathbf{F}_2$, ... (rys. 3.2), tworzących z kierunkiem przesunięcia odpowiednio kąty $\theta_1, \theta_2, \ldots$, to łączna praca wszystkich sił może być wyrażona jako

$$W = (F_1\cos\theta_1 + F_2\cos\theta_2 + \ldots)s = (F_{t_1} + F_{t_2} + \ldots)s,$$



Rys. 3.2

gdzie $F_{t_1}, F_{t_2}, \ldots$ są odpowiednio rzutami sił $\mathbf{F}_1, \mathbf{F}_2, \ldots$ na kierunek $\mathbf{s}$. Suma tych rzutów równa się rzutowi F_t siły wypadkowej $\mathbf{F} = \mathbf{F}_1 + \mathbf{F}_2 + \ldots$, czyli

$$F_{t_1} + F_{t_2} + \ldots = F_t.$$

A zatem i w tym przypadku całkowita praca wszystkich sił jest równa

$$W = F_t s.$$

Obliczanie pracy nieco się komplikuje, gdy ciało pod działaniem siły $\mathbf{F}$ zmiennej co do wartości posuwa się po torze prostoliniowym, z którym siła $\mathbf{F}$ stale tworzy kąt θ. Na przykład niech ciało przesuwa się wzdłuż osi x od położenia początkowego x_1 do położenia końcowego x_2.

Tym razem pracę obliczamy w ten sposób, że całkowitą drogę $x_2 - x_1$ dzielimy na bardzo wiele odcinków tak małych, aby siłę na nie działającą można było uważać za stałą. Praca całkowita równa się sumie prac elementarnych wykonanych na poszcze-

gólnych małych odcinkach drogi. Jeśli liczba odcinków jest n, a długość każdego z nich Δx, to

$$W = \cos\theta \sum_{i=1}^{n} F_i \Delta x_i = \sum_{i=1}^{n} F_{t_i} \Delta x_i.$$

Wartość tej sumy rozciągnięta na nieskończenie wiele zmierzających do zera elementów drogi sięgającej od x_1 do x_2 sprowadza się do całki oznaczonej

$$W = \int_{x_1}^{x_2} F_t dx. \tag{3.3}$$

Przykłady obliczania pracy w przypadku siły zmiennej co do wartości rozpatrujemy m.in. w § 4.6.

Najogólniejszy przypadek obliczania pracy dotyczy działania siły zmiennej co do wartości i kierunku, działającej wzdłuż drogi krzywoliniowej. Tym razem obliczanie pracy sprowadza się do obliczania całki krzywoliniowej. Tego przypadku szerzej nie będziemy omawiali.

Po tych rozważaniach o charakterze definicyjnym rozpatrzmy konkretny przykład prac wykonanych przez różne siły podczas ruchu ciała. Niech np. samochód przy pewnej określonej sile napędowej porusza się ruchem jednostajnym wzdłuż poziomej drogi s na prostoliniowym torze. Z pierwszej zasady dynamiki wynika, że wypadkowa siła działająca na samochód w tych warunkach równa się zeru. A zatem i całkowita praca W związana z ruchem samochodu równa się zeru. Ale na pracę całkowitą składa się praca W_1 siły napędowej F_1 oraz praca W_2 siły wypadkowej wszystkich sił oporu F_2 skierowanej przeciwnie w stosunku do prędkości. A zatem

$$W = W_1 + W_2 = 0,$$

skąd

$$W_1 = -W_2.$$

Uogólniając otrzymany wynik powiemy, że w przypadku ruchu jednostajnego ciała po torze poziomym dodatnia praca sił napędowych (tzw. praca włożona w ciało poruszające się) jest liczbowo równa ujemnej pracy sił oporowych (tzw. pracy wykonanej przez ciało poruszające się).

3.2. Moc

Z pojęciem pracy ściśle się wiąże pojęcie mocy. Jeśli w pewnym przedziale czasu Δt praca ΔW wykonywana jest równomiernie, to przez *moc* P rozumiemy stosunek pracy do czasu, w jakim została wykonana:

$$P = \frac{\Delta W}{\Delta t}. \tag{3.4}$$

Jeśli jednak praca wykonywana w równych, dowolnie krótkich odstępach czasu nie

jest stała, to posługujemy się pojęciem *mocy średniej* $\overline{P}$ i *mocy chwilowej* P zdefiniowanych odpowiednio równaniami:

$$\overline{P} = \frac{\Delta W}{\Delta t}, \quad P = \frac{dW}{dt}.$$

Jednostką mocy w układzie SI jest 1 wat (W). Zgodnie z (3.4)

$$1\ \mathrm{W} = 1\ \mathrm{J/s}^*.$$

Często też jest w użyciu jednostka zwana kilowatem (kW):

$$1\ \mathrm{kW} = 1000\ \mathrm{W}.$$

Pytanie: Jakimi krotnościami wata są: megawat (MW), miliwat (mW), mikrowat (μW)?

3.3. Energia

Pojęcie pracy ściśle się wiąże z pojęciem energii, pojęciem podstawowym dla całego przyrodoznawstwa i dla techniki.

Energia jest to zasób pracy, zmagazynowany w danym ciele lub układzie ciał, który może być zmniejszany (energia maleje) lub zwiększany (energia rośnie). Innymi słowy, wartość energii nie jest stała, charakterystyczna dla danego ciała (układu), lecz zależy od jego stanu. Do tej sprawy jeszcze powrócimy.

Energia może występować w dwóch podstawowych postaciach: jako energia związana z ruchem, czyli tzw. *energia kinetyczna*, oraz jako energia związana ze specjalnym położeniem elementów danego ciała (lub elementów wchodzących w skład układu ciał) względem siebie, czyli tzw. *energia potencjalna*. W zależności od charakteru rozpatrywanych zjawisk można rozróżniać energię mechaniczną, elektryczną, magnetyczną, jądrową itp.

W tej części kursu fizyki będzie nas interesowała energia mechaniczna, którą również można podzielić na energię kinetyczną i potencjalną. Zilustrujemy je przykładami.

Mówimy, że *ciało* ma *mechaniczną energię kinetyczną*, gdy dzięki *prędkości* swego ruchu zdolne jest do wykonywania pracy. Taki warunek spełnia np. wagon kolejowy poruszający się po szynach. Uderzając o jakąś przeszkodę może on ją przesunąć lub zgnieść, a więc wykonać pracę.

Jako przykład *mechanicznej energii potencjalnej* może służyć *energia potencjalna sprężysta*. Przypiszemy ją np. ściśniętej lub rozciągniętej sprężynie, która rozprężając się lub kurcząc może wykonać pracę. Inny rodzaj mechanicznej energii potencjalnej, a mianowicie energię potencjalną grawitacyjną ma np. ciało wzniesione na pewną wysokość nad ziemię: spadając na ziemię może ono podnieść inne ciało, a więc również wykonać pracę.

---

* Przejściowo dopuszczone do stosowania są m.in. następujące jednostki mocy, nie należące do układu SI:

$$1\ \mathrm{erg/s} = 10^{-7}\ \mathrm{W},$$
$$1\ \mathrm{kGm/s} = 9{,}80665\ \mathrm{W},$$
$$1\ \mathrm{KM} = 75\ \mathrm{kGm/s} \simeq 736\ \mathrm{W}.$$

Definicja energii podana wyżej jest raczej poglądowa, nie jest tak dokładna jak definicje innych rozważanych dotychczas wielkości fizycznych. Wiąże się to z tym, że *nie wyznaczamy całkowitej energii posiadanej przez ciało (czy układ ciał), a tylko jej zmiany — przyrosty lub ubytki — związane z przejściem ciała (układu) od określonego stanu początkowego do stanu końcowego.* Najczęściej stanowi początkowemu umownie przypisuje się pewną określoną wartość energii E_0 (E_0 może mieć np. umowną wartość zerową). W przypadku obliczania mechanicznej energii kinetycznej umownie określonym stanem początkowym (zerowym) jest stan spoczynku ciała ($v = 0$) w danym układzie odniesienia. W przypadku różnych rodzajów energii potencjalnych stan początkowy może być wybierany różnie: może to być np. stan zupełnie swobodnej sprężyny, nie napiętego łuku, lokalizacja ciała na powierzchni Ziemi ($h = 0$) itp.

3.3.1. *Energia kinetyczna.* Od rozważań ogólnych związanych z pojęciem energii przejdźmy do rozważań związanych z energią kinetyczną.

Do wzoru na energię kinetyczną doprowadza następujące rozumowanie. Na ciało o masie m, poruszające się z prędkością $\mathbf{v}_0$, zaczyna w pewnej chwili ($t = 0$) działać siła $\mathbf{F} = \text{const}$, skierowana zgodnie z kierunkiem prędkości. Siłę $\mathbf{F}$ traktujemy jako jedyną siłę działającą na poruszające się ciało: może to być albo rzeczywiście jedyna siła (np. siła ciężkości działająca na ciało spadające w próżni), albo wypadkowa wszystkich sił działających na dane ciało. Siła $\mathbf{F}$, działając w czasie t na pewnej drodze s, wywołuje ruch jednostajnie przyspieszony i wykonuje pracę W:

$$W = Fs = F(v_0 t + \tfrac{1}{2}at^2). \tag{3.5}$$

Praca ta zostaje zmagazynowana w ciele pod postacią przyrostu energii kinetycznej:

$$W = \Delta E_k = E_k^{(1)} - E_k^{(0)}, \tag{3.6}$$

gdzie $E_k^{(1)}$ i $E_k^{(0)}$ oznaczają odpowiednio energię kinetyczną końcową i początkową ciała.

Wyznaczając wartość włożonej pracy W w zależności od masy ciała i jego prędkości początkowej i końcowej znajdziemy wyrażenia na energię kinetyczną. Podstawiając do (3.5) $a = F/m$ otrzymujemy

$$W = Fv_0 t + \frac{F^2 t^2}{2m}.$$

Uwzględniając zależność między popędem siły i przyrostem pędu masy

$$Ft = mv_1 - mv_0,$$

gdzie v_1 jest prędkością końcową ciała po t sekundach, znajdujemy

$$W = mv_1 v_0 - mv_0^2 + \frac{m^2 v_1^2 - 2m^2 v_1 v_0 + m^2 v_0^2}{2m},$$

skąd ostatecznie

$$W = \frac{mv_1^2}{2} - \frac{mv_0^2}{2}. \tag{3.7}$$

Porównując (3.7) z (3.6) otrzymujemy

$$E_k^{(1)} = \frac{mv_1^2}{2} \quad \text{oraz} \quad E_k^{(0)} = \frac{mv_0^2}{2}.$$

Ogólnie zatem *energię kinetyczną ciała o masie m i prędkości v wyraża połowa iloczynu masy ciała i kwadratu jego prędkości.*

Gdy w chwili przyłożenia siły **F** = const ciało rozważane jest w spoczynku ($\mathbf{v}_0 = 0$), to stan początkowy energii kinetycznej jest zerowy i praca siły **F** zużywa się na nadanie ciału przyrostu energii kinetycznej równego po prostu końcowej wartości $E_k^{(1)}$:

$$W = \Delta E_k = E_k^{(1)} = \frac{mv_1^2}{2}.$$

Gdy przyłożona siła **F** = const jest skierowana przeciwnie do prędkości $\mathbf{v}_0$, wywołany przez nią ruch jest jednostajnie opóźniony. Ujemna praca tej siły powoduje ubytek energii kinetycznej. Tym razem kosztem energii kinetycznej zostaje wykonana praca pokonania siły oporu. Tak np., jeżeli założymy, że meteor, mający w chwili uderzenia o Ziemię prędkość $\mathbf{v}_0$, pokonuje przy zagłębianiu się stałą siłę oporu $\mathbf{F}_{op}$, to przebędzie on w Ziemi drogę s spełniającą równanie

$$F_{op}s = \frac{mv_0^2}{2}.$$

Gdy wypadkowa siła działająca na ciało poruszające się z prędkością **v** równa się zeru, ruch ciała jest jednostajny i energia kinetyczna zachowuje wartość stałą (por. przykład z samochodem w § 3.1).

W dotychczasowych rozważaniach braliśmy pod uwagę wypadkową siłę stałą, działającą wzdłuż prostoliniowej drogi. Oba te założenia nie są konieczne. Przy sile zmiennej i torze krzywoliniowym dzielimy całą drogę na bardzo wiele odcinków tak małych, aby każdy z nich można było uważać za prostoliniowy, a siłę na nim działającą za stałą. Praca wykonana na każdym z tych odcinków wywołuje pewien przyrost energii kinetycznej. Suma prac na wszystkich odcinkach daje sumaryczny przyrost energii kinetycznej badanego ciała.

Chyba nie wymaga dodatkowego podkreślenia fakt, że energia kinetyczna ciała ma wartość liczbową zależną od układu odniesienia. Wszak $E_k = f(v)$, a prędkość zależy od układu odniesienia. Tak np. pasażer siedzący w poruszającym się wagonie ma E_k równą zeru względem pociągu, lecz nierówną zeru w odniesieniu do Ziemi.

3.3.2. *Energia potencjalna mechaniczna.* Jak już wspominaliśmy w § 3.3 energia mechaniczna może być zmagazynowana w ciele (lub w układzie ciał) nie tylko pod postacią energii kinetycznej, lecz także pod postacią energii potencjalnej.

Tak np. podnosząc ciało w próżni ruchem jednostajnym na wysokość h nie wywołujemy przyrostu energii kinetycznej, gdyż wypadkowa siła działania naszych mięśni i siły ciężkości równa się zeru. Całkowita praca obu tych sił (jednej dodatniej, drugiej ujemnej) równa się zeru. Dodatnia praca naszych mięśni nie jest jednak „zmarnowana" — jest ona zmagazynowana w ciele podniesionym na wysokość h nad powierzchnię Ziemi i może być zwrócona przy spadku ciała na ziemię. Podobnie może być odzyskana praca włożona na ściśnięcie lub rozciągnięcie sprężyny, jak już o tym mówiliśmy. Praca włożona nie będzie jednak zwrócona, jeśli zostanie zużyta na pokonanie tarcia. Przesunięcie ciała na drodze poziomej s ruchem jednostajnym wymaga zasto-

sowania siły napędowej, równej sile tarcia. Podobnie jak w pierwszym przypadku całkowita praca obu tych sił równa się zeru. Tym razem jednak dodatnia praca siły napędowej nie zostaje zmagazynowana w ciele: powrót ciała do położenia początkowego wymaga znów zastosowania siły napędowej, siła tarcia bowiem znowu przeszkadza ruchowi. Praca siły napędowej zostaje rozproszona w otoczeniu pod postacią energii cieplnej.

W omówionych przykładach siła napędowa pokonywała kolejno siłę ciężkości, siłę sprężystą i siłę tarcia. Pierwsze dwie siły zaliczamy do tzw. *sił potencjalnych* (*zachowawczych*), trzecią — do *sił rozpraszających*. *Praca pokonania sił potencjalnych zostaje zmagazynowana w ciele pod postacią energii potencjalnej, natomiast praca pokonania sił rozpraszających zamienia się na energię cieplną i rozprasza w otoczeniu*. Dokładniejsze rozróżnienie sił zachowawczych i rozpraszających podamy niżej.

W związku z różnymi postaciami energii potencjalnej mechanicznej* wzory wyrażające ją zależą od rodzaju działających sił zachowawczych. Będziemy je wyprowadzali w odpowiednich rozdziałach (np. 19). Obecnie poprzestaniemy na wyprowadzeniu wzoru na energię potencjalną grawitacyjną, wynikającą z istnienia siły przyciągania ziemskiego.

Załóżmy, że ciało o masie m znajduje się na pewnym poziomie początkowym, któremu przypisujemy umownie zerową wartość energii potencjalnej. Energia potencjalna E_p związana z podniesieniem ruchem jednostajnym tego ciała na wysokość h ponad poziom zerowy powstaje kosztem pracy pokonania siły ciężkości (opór powietrza zaniedbujemy). A zatem

$$E_p = W = mgh. \tag{3.8}$$

Gdyby ciało znajdując się na poziomie początkowym już posiadało pewną energię potencjalną $E_p^{(0)} = \text{const}$, to wspomniana praca zużyłaby się na wywołanie przyrostu energii potencjalnej od wartości początkowej $E_p^{(0)}$ do końcowej $E_p^{(1)}$:

$$mgh = E_p^{(1)} - E_p^{(0)}.$$

A zatem

$$E_p^{(1)} = mgh + E_p^{(0)} \quad \text{lub} \quad E_p^{(1)} = mgh + \text{const}.$$

Z tych wzorów widać wyraźnie, że *praca włożona na przesunięcie ciała z poziomu początkowego do końcowego nie wyznacza całkowitego zasobu końcowej energii potencjalnej, a tylko jej nadwyżkę ponad stałą wartość $E_p^{(0)}$, określającą energię potencjalną na poziomie początkowym*. Zazwyczaj poziom początkowy przyjmujemy za zerowy i energię potencjalną grawitacyjną względem tego poziomu wyznaczamy ze wzoru (3.8).

Bliższe zbadanie energii potencjalnej grawitacyjnej prowadzi do wniosku, że wartość energii potencjalnej ciała wzniesionego na wysokość h ponad poziom zerowy nie zależy od drogi (rys. 3.3), wzdłuż której zostało ono podniesione, przy założeniu braku tarcia. Z różnych dróg przedstawionych na rys. 3.3 wybierzmy drogę pierwszą — pionową i drogę drugą — ukośną (np. po równi pochyłej bez tarcia); przedstawmy je na rys. 3.4 i dla nich przeprowadźmy dowód.

---

* Przypominamy, że energia potencjalna nie ogranicza się do energii potencjalnej mechanicznej. Na przykład o energii potencjalnej elektrycznej i magnetycznej mówimy w rozdz. 19, 21, 30.

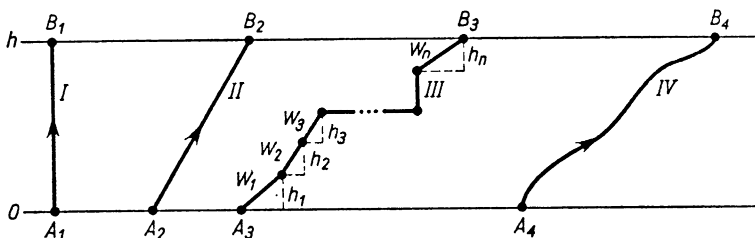
Praca wzniesienia ciała o masie m na wysokość h wzdłuż drogi BC równa się, jak już wiemy,

$$W = mgh.$$

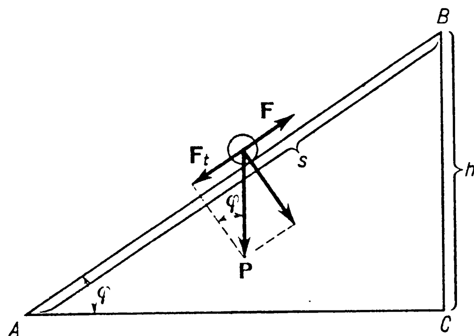
Wciąganie ciała po równi pochyłej ruchem jednostajnym wymaga zastosowania siły **F** równoważącej składową $\mathbf{F}_t$ siły ciężkości styczną do powierzchni równi. Ale

$$F_t = mg\sin\varphi,$$

gdzie φ oznacza kąt nachylenia równi względem poziomu.



Rys. 3.3



Rys. 3.4

Praca siły wciągającej ciało na drodze $AB = s$ równa się

$$W' = mg\sin\varphi \cdot s.$$

Ale

$$s\sin\varphi = h.$$

A zatem praca wciągania ciała po równi pochyłej na wysokość h

$$W' = mgh = W$$

jest taka sama jak praca przy pionowym podnoszeniu ciała na tę samą wysokość (nie zależy od kąta φ).

Po tym wstępnym dowodzie powróćmy do rys. 3.3. Droga łamana III składa się z odcinków poziomych, pionowych i ukośnych. Praca przesuwania ciała ruchem jedno-

stajnym bez tarcia na odcinkach poziomych równa się zeru. Na odcinkach pionowych i ukośnych praca ta zgodnie z podanym wyżej dowodem równa się iloczynowi mg przez odpowiednią wysokość. Całkowita praca przesunięcia ciała z punktu A_3 do B_3 równa się sumie prac wykonanych na poszczególnych odcinkach linii łamanej, czyli

$$W = W_1 + W_2 + W_3 + \ldots + W_n,$$

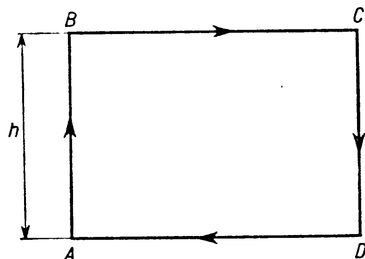
$$W = mg(h_1 + h_2 + h_3 + \ldots + h_n),$$

$$W = mgh.$$

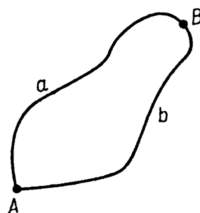
Przesunięcie wzdłuż toru krzywoliniowego IV można rozbić na bardzo wiele odcinków tak małych, że każdy z nich można uważać za prostoliniowy. Wtedy rozumowanie analogiczne jak w przypadku drogi III doprowadza do wyrażenia

$$W = mgh.$$

Stwierdziliśmy zatem fakt o pierwszorzędnym znaczeniu: *praca pokonywania siły ciężkości (a więc i praca siły ciężkości) nie zależy od drogi, lecz tylko od stanu początkowego i końcowego.* Można wykazać, że jest to *słuszne dla wszystkich sił potencjalnych.*



Rys. 3.5



Rys. 3.6

Na przykładzie siły ciężkości można też sprawdzić drugą właściwość sił potencjalnych, a mianowicie, *że praca wykonana przez siłę potencjalną na drodze zamkniętej równa się zeru.* Niech np. tor zamknięty ma postać przedstawioną na rys. 3.5. Ciało o masie m przesuwa się wzdłuż drogi $ABCDA$. Na odcinku AB praca siły ciężkości $W_{AB} = -mgh$, na odcinkach BC i DA prace są równe zeru, a na odcinku CD praca $W_{CD} = +mgh$. A zatem $\sum W = 0$. Oczywiście można byłoby tor zamknięty przedstawić w postaci dowolnej linii krzywej. Wtedy należałoby podzielić go na nieskończenie wiele małych odcinków (w granicy prostoliniowych) i rozważać na nich prace siły ciężkości. Wynik ostateczny byłby ten sam: praca wykonana przez siłę ciężkości na drodze zamkniętej równa się zeru. Uogólnienie tego wyniku obejmuje wszystkie siły potencjalne.

Obie omówione właściwości sił potencjalnych (zachowawczych) — niezależność pracy od kształtu drogi łączącej punkt początkowy i końcowy oraz praca na drodze zamkniętej równa zeru — nie są właściwościami niezależnymi od siebie. Jedną z nich można uważać za wniosek wypływający z drugiej. Każda z nich jest tak charakterystyczna dla siły zachowawczej, że może służyć jako podstawa do zaliczenia badanej siły do kategorii sił zachowawczych.

Pytanie. Pod działaniem pewnej siły zachowawczej odbywa się przesunięcie ciała od punktu A (rys. 3.6) wzdłuż drogi a do punktu B, a następnie powrót wzdłuż drogi b do punktu A. W jaki sposób można udowodnić, że praca przesunięcia ciała od A do B nie zależy od kształtu drogi, jeśli wiadomo, że praca na drodze zamkniętej $AaBbA$ równa się zeru?

3.4. Układy zachowawcze i rozpraszające. Zasada zachowania energii mechanicznej

Ostatnie nasze rozważania dotyczyły takich układów ciał, w których działały wyłącznie siły zachowawcze. *Takie układy nazywamy zachowawczymi.* Układy zachowawcze w ścisłym znaczeniu nie istnieją w przyrodzie. Układ zachowawczy stanowiłaby np. idealnie sprężysta kula odbijająca się w próżni od idealnie sprężystej podstawy. W przybliżeniu układem zachowawczym byłoby wahadło odbywające ruch w próżni z minimalnym tarciem w punkcie zawieszenia. Prawie idealnie zachowawczy jest układ słoneczny.

W układach zachowawczych odosobnionych (tzn. nie poddanych oddziaływaniom z zewnątrz) obowiązuje zasada zachowania energii mechanicznej. Według tej zasady w *układzie zachowawczym odosobnionym całkowita energia mechaniczna E, równa sumie energii kinetycznej i potencjalnej, jest wielkością stałą, tzn. niezmienną w czasie:*

$$E = E_k + E_p = \text{const}.$$

Treść tej zasady sprowadza się do tego, że jeśli w układzie zachowawczym odosobnionym ciało odbywa ruch, to każdej zmianie jego energii kinetycznej towarzyszy równa co do wartości, lecz przeciwna co do znaku zmiana energii potencjalnej.

Układy, z jakimi mamy do czynienia na Ziemi, to *układy rozpraszające*, gdyż występuje w nich zwykle tarcie, a więc siła rozpraszająca. A zatem w przypadku przemiany pracy na energię w warunkach ziemskich kosztem pracy siły napędowej zmienia się zasób energii mechanicznej kinetycznej i potencjalnej i dodatkowo pojawiają się nowe postaci energii. Zilustrujemy to wykorzystując raz jeszcze przykład z samochodem. Tym razem niech on wjeżdża ruchem przyspieszonym na górę po nawierzchni drogi o pewnym tarciu. Kosztem pracy siły napędowej rośnie energia kinetyczna (wzrost prędkości), rośnie też energia potencjalna grawitacyjna (wzrost wysokości wzniesienia) oraz pojawia się energia cieplna związana z pokonywaniem siły tarcia.

3.5. Zasada zachowania energii

Mimo że w niniejszym rozdziale zajmowaliśmy się przede wszystkim energią mechaniczną, nie może w nim zabraknąć jednej z najważniejszych zasad całego przyrodoznawstwa, a mianowicie zasady zachowania energii. Dotyczy ona wszelkich możliwych odmian energii. *Według tej zasady w układzie odosobnionym od zewnętrznego otoczenia w ten sposób, że energia w żadnej postaci nie przenika do niego z zewnątrz ani nie uchodzi z niego na zewnątrz, całkowita wartość energii pozostaje niezmienna: mogą w nim tylko zachodzić*

*przemiany energetyczne jednej postaci energii w inną**. Energia nie może być ani stwarzana, ani niszczona.

Wnioskiem z zasady zachowania energii, potwierdzonej w niezliczonych badaniach, jest niemożliwość zbudowania urządzenia zwanego *perpetuum mobile*, które pracowałoby bez zasilania energią z zewnątrz i bez zmniejszania zasobu energii własnej.

Pytania i zadania

1. Jak brzmi ścisła definicja pracy i jej jednostki w układzie SI?

2. Czy może się zdarzyć, że praca równa się zeru, mimo że $F \neq 0$ i $s \neq 0$? Uzasadnić odpowiedź i podać przykłady.

3. Podać definicję mocy i jej jednostki w układzie SI.

4. Jakie rozumowanie doprowadza do wzoru na energię kinetyczną?

5. Z jakim rodzajem sił związane jest pojęcie energii potencjalnej?

6. Omówić zasadnicze różnice między siłami potencjalnymi i rozpraszającymi.

7. Omówić ogólną metodę obliczania energii potencjalnej. Czy otrzymany wynik dokładnie określa zasób energii potencjalnej w stanie końcowym?

8. Czym się różni zasada zachowania energii mechanicznej od zasady zachowania energii?

9. Kuleczkę wahadła matematycznego o masie m wiszącą na nici o długości L odchylono tak, że nić tworzy kąt α z pionem. Obliczyć, z jaką prędkością mija swobodnie puszczona kuleczka położenie równowagi i jaka jest w tej chwili siła naprężająca nić?

Odp. $2\sin\frac{\alpha}{2}\sqrt{gL}$, $\quad mg+4mg\sin^2\frac{\alpha}{2}$.

10. Gdy samochód o masie m miał na torze poziomym prędkość v_0, włączono silnik i dzięki temu na drodze s prędkość wzrosła do wartości v_1. Przyjmując, że siły oporu wynoszą Q, obliczyć moc silnika.

Odp. $\frac{m(v_1^2-v_0^2)(v_1+v_0)}{4s}+\frac{Q(v_1+v_0)}{2}$.

11. Obliczyć energię potencjalną ciała o masie m, wykonującego pod działaniem sił sprężystych ruch opisany zależnością $x = A\sin\omega t$ w chwili, gdy jego wychylenie od położenia równowagi ($x_0 = 0$) wynosi x_1.

Odp. $\frac{m\omega^2 x_1^2}{2}$

* O równoważności masy i energii mówimy w rozdz. 8.

ROZDZIAŁ 4

Ciążenie powszechne

4.1. Prawo powszechnego ciążenia

Prawo powszechnego ciążenia, zwane również prawem grawitacji, zostało sformułowane przez Newtona. Ściśle biorąc, w podanej niżej postaci odnosi się ono do punktów materialnych, tzn. do ciał o rozmiarach takich, że można je zaniedbać wobec wzajemnej odległości ciał.

Dwa punkty materialne o masach m_1 i m_2 przyciągają się wzajemnie siłą proporcjonalną do iloczynu ich mas i odwrotnie proporcjonalną do kwadratu ich odległości:

$$F = G\frac{m_1 m_2}{r^2}. \tag{4.1}$$

Współczynnik proporcjonalności G nosi nazwę *stałej grawitacyjnej*. Wartość tej stałej zależy tylko od układu jednostek. W układzie SI stała grawitacyjna równa się $6{,}67 \cdot 10^{-11}$ $N \cdot m^2/kg^2$. Jest to jedna z podstawowych fizycznych stałych uniwersalnych.

Stała grawitacyjna wyraża liczbowo siłę wzajemnego przyciągania dwóch punktów materialnych o masie 1 kilograma każdy z odległości 1 metra.

Stała grawitacyjna jako jedna z podstawowych stałych fizycznych była wyznaczana w różnych okresach i różnymi metodami. Pierwszego pomiaru dokonał Cavendish stosując czułą wagę skręceń. Z kursu szkoły średniej znana jest metoda Jolly'ego. Zwykle przy tego rodzaju pomiarach stosowane są ciała pomocnicze o bardzo dużych masach, rzędu setek ton. Jest to konieczne, gdyż w związku z bardzo małą wartością stałej grawitacyjnej siły ciążenia między ciałami o niewielkich masach są bardzo nieznaczne.

Przy obliczaniu siły wzajemnego przyciągania ciał o skończonych rozmiarach traktujemy je jako zbiorowisko punktów materialnych i stosując rachunek całkowy obliczamy wypadkowe oddziaływanie wszystkich punktów materialnych. Postępując w ten sposób można np. stwierdzić, że ciała kuliste jednorodne lub ciała kuliste złożone z warstw kulistych jednorodnych przyciągają się tak, jakby ich masy były skupione w ich środkach.

4.2. Stosowalność prawa powszechnego ciążenia do ciał niebieskich

Wiemy już, że stała grawitacyjna jest stałą uniwersalną zależną tylko od wybranego układu jednostek, natomiast niezależną od tego, czy przyciągające się masy są w spoczynku, czy w ruchu, są małe czy duże i w jakim znajdują się ośrodku.

Spróbujmy idąc śladami Newtona przeanalizować, czy prawo powszechnego ciążenia odnosi się do Księżyca i Ziemi. Ciała te można traktować jako punkty materialne: odległość Ziemia — Księżyc wynosi ok. 60 promieni Ziemi (ok. 385 000 km), promień Ziemi — ok. 6400 km, a promień Księżyca — ok. $0{,}273 R_{Ziemi}$, czyli ok. 1700 km.

Gdyby ciało o masie Księżyca znalazło się w odległości jednego promienia Ziemi od jej środka, czyli na jej powierzchni, to spadałoby ku Ziemi z przyspieszeniem ziemskim $g = 9{,}81$ m/s^2. To samo ciało odsunięte na odległość $60R_Z$ podlega — według prawa grawitacji — sile 60^2 razy mniejszej, a więc powinno spadać ku Ziemi z przyspieszeniem

$$\frac{9{,}81}{3600} \text{ m/s}^2 = 0{,}00273 \text{ m/s}^2.$$

Jeśli zatem Księżyc podlega prawu grawitacji, to w jego ruchu dokoła Ziemi (w przybliżeniu po okręgu) powinno występować przyspieszenie dośrodkowe równe 0,00273 m/s^2. Ale przyspieszenie dośrodkowe Księżyca można obliczyć znając okres T jego obiegu dokoła Ziemi równy 2 360 591,5 sekund:

$$a = \omega^2 r = \frac{4\pi^2}{T^2} 60R_Z \approx 0{,}00271 \text{ m/s}^2.$$

Zgodność wyników jest dostateczna do wyciągnięcia wniosku, że prawo powszechnego ciążenia stosuje się do Księżyca i Ziemi. Uogólniamy ten wniosek na pozostałe ciała niebieskie.

P y t a n i e: Jakimi wzorami można wyrazić siłę dośrodkową powodującą ruch Księżyca dokoła Ziemi?

4.3. Prawa Keplera

Wielkim sukcesem Newtona było matematyczne wyprowadzenie praw Keplera na podstawie prawa powszechnego ciążenia i zasad dynamiki. *Prawa Keplera* potwierdzające heliocentryczną teorię Kopernika były prawami wysnutymi z bogatego materiału obserwacyjnego dotyczącego ruchów planet, pochodzącego od różnych astronomów i zestawionego przez Tychona de Brahe. Były to zatem wnioski z obserwacji, które dopiero dzięki pracom Newtona zyskały uzasadnienie teoretyczne.

Według *pierwszego prawa Keplera wszystkie planety poruszają się po torach eliptycznych, w których wspólnym ognisku znajduje się Słońce.*

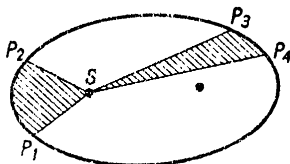
Drugie prawo określa prędkość ruchu poszczególnych planet na ich orbitach: *planeta porusza się w ten sposób, że pola zakreślane w równych czasach przez promień wodzący poprowadzony od Słońca S do planety P są sobie równe* (rys. 4.1). Wnioskiem z tego prawa jest stwierdzenie niejednostajności ruchu planet: gdy planeta znajduje się bliżej Słońca, to większa jest jej prędkość liniowa.

Według trzeciego prawa kwadraty okresów obiegów poszczególnych planet dokoła Słońca są proporcjonalne do sześcianów ich średnich odległości od Słońca:

$$T_1^2 : T_2^2 = b_1^3 : b_2^3.$$

Średnią odległością b planety od Słońca jest połowa wielkiej osi elipsy.

Potwierdzeniem stosowalności prawa powszechnego ciążenia do ciał niebieskich jest też występowanie perturbacji w ruchach planet. *Perturbacje* są to odchylenia od właściwych planetom torów eliptycznych. Gdyby planety podlegały tylko działaniu przyciągającemu Słońca, to ich tory zgodnie z I prawem Keplera byłyby eliptyczne. Zakłócenia w tych ruchach wywołane są działaniem grawitacyjnym pozostałych planet. Z wielkości i rodzaju zakłóceń, występujących w ruchu badanej planety, można wnioskować o rozmieszczeniu i masie innych planet w jej sąsiedztwie. Takie przeliczenia doprowadziły do wykrycia planet Neptuna i Plutona.



Rys. 4.1

4.4. Ciężar ciał na Ziemi

Prawo powszechnego ciążenia stosujemy również do Ziemi, uważając ją w przybliżeniu za ciało złożone z warstw kulistych jednorodnych. Wtedy ciężar P punktu materialnego na powierzchni Ziemi można traktować jako siłę spełniającą prawo powszechnego ciążenia:

$$P = G\frac{m_Z m}{R_Z^2},$$

gdzie m_Z oznacza masę Ziemi, R_Z — promień Ziemi, m — masę ciała.

Z drugiej zasady dynamiki

$$P = mg.$$

Z porównania obu wzorów wynika, że

$$g = G\frac{m_Z}{R_Z^2}, \tag{4.2}$$

czyli przyspieszenie ziemskie nie zależy od masy ciała. Zależy natomiast od odległości od środka Ziemi. Na wysokości h nad jej powierzchnią

$$g = G\frac{m_Z}{(R_Z+h)^2}.$$

Na ogół wzniesienia h nad powierzchnię Ziemi są małe w porównaniu z promieniem Ziemi, a więc poprawka g na wysokość wzniesienia jest niewielka.

Rozpatrzmy obecnie zmiany ciężaru P, wynikające z zagłębiania się ciała pod powierzchnię Ziemi (rys. 4.2). Niech badane ciało (punkt materialny) znajduje się w odległości r od środka Ziemi. Można udowodnić, że wypadkowe przyciąganie masy badanej m przez warstwę kulistą o promieniu wewnętrznym r i zewnętrznym R (zakreskowaną na rysunku) równa się zeru. Pozostaje zatem tylko przyciąganie kuli o promieniu r.

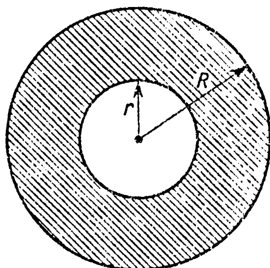
Masa m_k tej kuli (przy założeniu, że jest ona jednorodna) równa się iloczynowi objętości i gęstości ϱ:

$$m_k = \tfrac{4}{3}\pi r^3 \varrho.$$

Siła przyciągania F, czyli ciężar ciała P w odległości r od środka Ziemi

$$P = G\frac{\frac{4}{3}\pi r^3 \varrho m}{r^2}, \qquad P = G \cdot \tfrac{4}{3}\pi r \varrho m,$$

czyli P jest wprost proporcjonalne do r. Przy zagłębianiu się pod powierzchnię Ziemi ciężar ciała jest proporcjonalny do odległości r od środka Ziemi. W środku Ziemi ciężar ciał równa się zeru*.



Rys. 4.2

Prawo powszechnego ciążenia umożliwia m. in. obliczenie masy Ziemi. Z równania

$$mg = G\frac{m_Z m}{R_Z^2}$$

wynika, że

$$m_Z = \frac{gR_Z^2}{G}, \qquad \text{czyli} \qquad m_Z \approx 6 \cdot 10^{24}\ \text{kg}.$$

4.5. Zmienność wartości przyspieszenia ziemskiego na Ziemi

Na wartość przyspieszenia w różnych punktach Ziemi wpływają: a) kształt Ziemi, b) ruch obrotowy Ziemi dokoła własnej osi, c) niejednorodność budowy Ziemi.

Ziemia ma swoisty kształt tzw. *geoidy*, przypominający elipsoidę obrotową spłaszczoną od strony biegunów. Dzięki temu wartość g rośnie w miarę wzrostu szerokości geograficznej. Każde ciało znajdujące się na Ziemi uczestniczy w jej ruchu obrotowym dokoła osi. Potrzebna do utrzymania ciała w tym ruchu siła dośrodkowa jest proporcjo-

* Trzeba tu podkreślić, że aczkolwiek ciężar jest siłą, to jednak wyrażany jest zazwyczaj nie w niutonach, lecz w specjalnych, nie należących do układu SI, jednostkach ciężarowych, z których wymienimy 2 najpospoliciej używane. Są to Gram-siła (G) zwany inaczej pondem (p) oraz Kilogram-siła (kG) zwany też kilopondem (kp). Jednostka Gram-siła jest to siła, z jaką Ziemia przyciąga masę 1 grama w tym miejscu, gdzie przyspieszenie ziemskie ma wartość tzw. normalną, równą $g_n = 9{,}80665\ \text{m/s}^2$. A zatem

$$1\ \text{G} = 1\ \text{p} = 0{,}009\,806\,65\ \text{N}, \qquad 1\ \text{kG} = 1\ \text{kp} = 9{,}806\,65\ \text{N}.$$

nalna (przy stałym m i ω) do promienia zakreślanego koła, a więc największa na równiku, a najmniejsza (zerowa) na biegunie. Siła dośrodkowa powstaje kosztem części siły przyciągania ziemskiego. Stąd wniosek, że ruch obrotowy Ziemi dokoła własnej osi wpływa na wartość g w ten sposób, że rośnie ono stopniowo w miarę przesuwania się wzdłuż południka od równika do bieguna (a więc także rośnie ze wzrostem szerokości geograficznej). Oba wymienione czynniki powodują, że g jest na biegunie większe niż na równiku. Niejednorodność struktury Ziemi, jak również ukształtowanie powierzchni Ziemi powodują niewielkie lokalne wahania wartości g.

Przyspieszenie ziemskie wyznaczamy doświadczalnie badając ruchy odbywające się pod działaniem siły ciężkości. Najdokładniejsze metody wiążą się z ruchami okresowymi, m.in. z ruchem wahadła matematycznego i fizycznego. Metody te omówimy w p. 10.2 i 10.3.

4.6. Pole grawitacyjne

Oddziaływania grawitacyjne między punktami materialnymi wykorzystamy do bliższego omówienia właściwości tzw. pola grawitacyjnego.

Wyobraźmy sobie pewien obszar absolutnie pusty, w którym znalazła się jedna cząstka materialna. Wobec zupełnej pustki w otoczeniu na tę cząstkę nie działa żadna siła. Jeśli jednak w jej pobliżu znajdzie się druga cząstka, to ona już będzie podlegała pewnemu działaniu. Można wtedy mówić o bezpośrednim oddziaływaniu na odległość cząstki I na cząstkę II, tak jak to robiliśmy dotychczas, lecz można też przyjąć inny punkt widzenia. Można uważać, że obecność cząstki I w pewnym obszarze wpłynęła na właściwości tego obszaru, a mianowicie wytworzyła w tym obszarze pole grawitacyjne objawiające się w ten sposób, że dowolny punkt materialny (np. cząstka II) wniesiony do tego obszaru podlega działaniu pewnej siły (grawitacyjnej) i zyskuje pewną energię (potencjalną). W tym ujęciu nie ma bezpośredniego działania między cząstką I i II; istnieje pewien pośrednik, którym jest wytworzone pole: cząstka I wytwarza pole w swym otoczeniu, a istniejące pole działa na cząstkę II.

Pole grawitacyjne nie jest jedynym polem, z którym mamy do czynienia w fizyce. Wystarczy przypomnieć pole elektrostatyczne znane z kursu szkolnego. Ładunek elektryczny umieszczony w pewnym obszarze wytwarza dokoła siebie pole objawiające się w ten sposób, że następny ładunek umieszczony w pobliżu podlega działaniu elektrostatycznemu. Pojęcie pól jest obecnie szeroko stosowane. Przejdziemy obecnie do zdefiniowania dwóch podstawowych wielkości charakteryzujących poszczególne punkty pola grawitacyjnego, a mianowicie wielkości wektorowej — natężenia pola oraz wielkości skalarnej — potencjału pola.

Opisując właściwości pola grawitacyjnego, wytworzonego przez punkt materialny o masie M, posługujemy się niewielką masą próbną m. Siła działająca na masę próbną umieszczoną w odległości r od M wyraża się wzorem

$$F = G\frac{Mm}{r^2} \tag{4.3}$$

lub w ujęciu wektorowym —

$$\mathbf{F} = -G\frac{Mm}{r^2}\frac{\mathbf{r}}{r} \tag{4.4}$$

(znak minus wiąże się z tym, że wektor **r** ma zwrot przeciwny niż wektor **F**). Siła **F** opisana wzorem (4.4) nie charakteryzuje w sposób jednoznaczny tego punktu pola, gdzie się znajduje masa próbna *m*, gdyż zależna jest od wartości tej próbnej masy. Można jednak obliczyć siłę, która by działała na masę jednostkową i w ten sposób przejść do wielkości zwanej *natężeniem pola grawitacyjnego*, oznaczanej literą **g**. Opierając się na tej definicji i na równaniu (4.4) można napisać następujące zależności:

$$\mathbf{g} = \frac{\mathbf{F}}{m} = -G\frac{M}{r^2}\frac{\mathbf{r}}{r}. \tag{4.5}$$

Warto podkreślić, że aczkolwiek wektor **g** liczbowo odpowiada sile działającej na jednostkową masę próbną umieszczoną w danym punkcie pola, to jednak wyraża się w innych jednostkach, a mianowicie w m/s^2, czyli w jednostkach przyspieszenia. Łatwo to sprawdzić. Wszak siłę działającą na swobodną masę w polu grawitacyjnym można przedstawić jako iloczyn tej masy i przyspieszenia **a**, z jakim masa *m* swobodnie spada w stronę masy *M*:

$$\mathbf{F} = m\mathbf{a},$$

a stąd

$$\mathbf{g} = \frac{\mathbf{F}}{m} = \mathbf{a}.$$

W przypadku pola grawitacyjnego Ziemi natężenie pola jest równe przyspieszeniu swobodnego spadania ku Ziemi, czyli przyspieszeniu ziemskiemu.

Wzory (4.3)–(4.5) są wzorami ogólnymi: dotyczą dowolnego pola grawitacyjnego. Rozpatrując pole grawitacyjne Słońca, Ziemi, Księżyca itp. trzeba zamiast *M* odpowiednio podstawiać m_S, m_Z, m_K itd.

Do charakterystyki danego punktu pola grawitacyjnego można też stosować wielkość zwaną potencjałem. *Potencjał pola* ściśle się wiąże z energią potencjalną, o której już była mowa w p. 3.3.2. Pamiętamy, że energię potencjalną można wyrazić jako pracę, którą trzeba wykonać przy przejściu od dowolnego stanu odniesienia do stanu końcowego. Rozważając energię potencjalną grawitacyjną ciał w pobliżu Ziemi wygodnie nam było przyjmować jako poziom odniesienia powierzchnię Ziemi. W ogólnych rozważaniach „polowych" postępuje się jednak inaczej. Jako punkt odniesienia przyjmuje się taki punkt, w którym siła działająca na masę próbną równa się zeru. W polu sił grawitacyjnych taki punkt będzie się znajdował w odległości nieskończonej od centrum przyciągania (podobnie postępuje się w elektrostatyce). A zatem energia potencjalna dowolnej masy w nieskończoności jest równa zeru. Praca siły zachowawczej grawitacyjnej $W_{\infty r}$ związana z przeniesieniem masy *m* z nieskończoności do punktu końcowego leżącego w odległości *r* od centrum przyciągania pociąga za sobą zmianę energii potencjalnej

$$\Delta E_p = E_{p_r} - E_{p\infty} = E_{p_r},$$

a zatem

$$E_{p_r} = W_{\infty r} = \int_{\infty}^{r} F dr = \int_{\infty}^{r} G\frac{Mm}{r^2}dr = -G\frac{Mm}{r}. \quad (4.6)$$

Znak minus w tym wzorze przypomina, że każdemu punktowi pola leżącemu w skończonej odległości r od centrum przyciągania odpowiada energia potencjalna mniejsza od $E_{p\infty}$ (którą przyjęliśmy za równą zeru), a więc energia ujemna. Najmniejsze energie potencjalne (tzn. największe wartości bezwzględne) charakteryzują punkty pola bliskie centrum przyciągania. W miarę oddalania się od centrum energia potencjalna rośnie (jej wartość bezwzględna maleje), osiągając wartość maksymalną równą zeru w nieskończoności. Takie zmiany energii potencjalnej są charakterystyczne dla sił przyciągających. Analogicznie np. przebiegają zmiany energii potencjalnej elektronu w polu dodatnio naładowanego jądra atomu.

Powróćmy raz jeszcze do wzoru (4.6). Widać z niego, że energia potencjalna, jaką posiada masa próbna m w danym punkcie pola grawitacyjnego, jest zależna od wartości masy próbnej. Obliczając wartość energii potencjalnej przypadającej na jednostkową masę próbną znajdujemy drugą wielkość charakteryzującą dany punkt pola grawitacyjnego, a mianowicie *potencjał* V_r:

$$V_r = \frac{E_{p_r}}{m} = -\frac{GM}{r}. \quad (4.7)$$

Łatwo sprawdzić, że w układzie SI jednostką potencjału grawitacyjnego jest J/kg (lub m^2/s^2).

Znów trzeba podkreślić, że wzory (4.6) i (4.7) mają znaczenie ogólne i po wstawieniu odpowiednich wartości na M mogą dotyczyć pola grawitacyjnego Słońca, Ziemi, Księżyca itp.

Potencjał grawitacyjny (podobnie jak dobrze znany potencjał elektryczny) jest wielkością skalarną. A zatem znalezienie potencjału dowolnego punktu pola grawitacyjnego wytworzonego przez szereg punktów materialnych M_1, M_2, M_3 itp. sprowadza się do znalezienia potencjałów V_1, V_2, V_3 itd. (odpowiadających poszczególnym centrom przyciągania) i ich arytmetycznego zsumowania. W odniesieniu do natężenia pola należałoby w tym przypadku zastosować sumowanie geometryczne.

4.7. Ruchy ciał w polu grawitacyjnym Ziemi

Z drugiej zasady dynamiki wynika, że pod działaniem stałej siły powstają ruchy jednostajnie zmienne. Rozpatrzmy jako przykład ruchy pod działaniem siły ciężkości: 1) swobodny spadek ciał, 2) rzut pionowy do góry, 3) rzut poziomy oraz 4) rzut ukośny.

1. *W swobodnym spadku ciał*, a więc w ruchu o prędkości początkowej $v_0 = 0$ i przyspieszeniu $a = g$, przebyta droga wyraża się wzorem

$$s = \frac{gt^2}{2}.$$

Po przekształceniu znajdujemy wyrażenie określające czas spadania w zależności od s:

$$t = \sqrt{\frac{2s}{g}}. \tag{4.8}$$

Oczywiście wzór ten jest słuszny i w tym przypadku, gdy s odpowiada całkowitej wysokości h, z jakiej ciało spada swobodnie. Prędkość końcowa ciała po przebyciu drogi h wynika ze wzoru

$$v = gt,$$

$$v = g\sqrt{\frac{2h}{g}} = \sqrt{2gh}. \tag{4.9}$$

2. *Rzut pionowy do góry* jako ruch jednostajnie opóźniony z opóźnieniem g wymaga nadania ciału prędkości początkowej v_0. Ta prędkość decyduje o maksymalnej wysokości rzutu h_{max} i o czasie wznoszenia się do góry t_{wzn}. Ruch do góry trwa do chwili, gdy prędkość końcowa staje się równa zeru, a zatem

$$0 = v_0 - gt_{wzn}.$$

Stąd

$$t_{wzn} = \frac{v_0}{g}. \tag{4.10}$$

Podstawiamy odpowiednie wartości do ogólnego wzoru na drogę w ruchu jednostajnie opóźnionym:

$$s = v_0 t - \frac{at^2}{2}$$

i znajdujemy

$$h_{max} = v_0 \frac{v_0}{g} - \frac{gv_0^2}{2g^2},$$

skąd

$$h_{max} = \frac{v_0^2}{2g}. \tag{4.11}$$

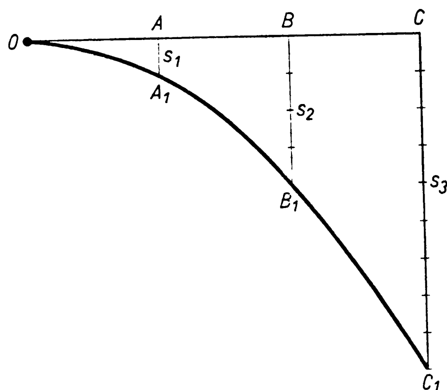
Elementarne rozważanie, które warto samodzielnie przeprowadzić, wykazuje, że czas wznoszenia się do wysokości h_{max} i czas swobodnego spadku z tej samej wysokości są równe.

P y t a n i e. Ciału nadano prędkość początkową v_0 skierowaną pionowo w dół. Jakimi wzorami wyrażają się w tym przypadku prędkość i przebyta droga jako funkcje czasu?

Dwa dalsze przypadki, 3 i 4, ruchów pod działaniem siły ciężkości są już bardziej skomplikowane, gdyż dotyczą ruchów złożonych. Ich badanie opiera się na *zasadzie niezależności ruchów* (por. § 1.6).

3. W przypadku *rzutu poziomego* nadajemy ciału prędkość początkową v_0 w kierunku poziomym. Gdyby ciało nie pozostawało pod działaniem żadnej siły, to zachowując tę prędkość poruszałoby się ruchem poziomym jednostajnym mijając po upływie kolej-

nych sekund punkty A, B, C (rys. 4.3). Siła ciężkości jednak działając na ciało powoduje jego równoczesne opadanie ruchem jednostajnie przyspieszonym z przyspieszeniem g. Drogi pionowe rosną proporcjonalnie do kwadratu czasu. Torem ciała jest parabola $OA_1B_1C_1$. Rzut poziomy dokładnie przeanalizowaliśmy jako przykład w p. 1.4.3.



Rys. 4.3

Zakładając, że układ odniesienia jest tak dobrany, że ruch jednostajny z prędkością v_0 odbywa się w kierunku osi x, a ruch jednostajnie zmienny z przyspieszeniem g w kierunku ujemnym osi y, otrzymaliśmy jako równanie toru wyrażenie

$$y = -\frac{g}{2v_0^2}x^2, \tag{4.12}$$

czyli równanie paraboli. Obliczmy *zasięg D rzutu poziomego*, gdy ciało w kierunku pionowym przesunie się o wysokość h. Do równania (4.12) wstawiamy $y = -h$, $x = D$ i otrzymujemy

$$-h = -\frac{g}{2v_0^2}D^2,$$

$$D = v_0\sqrt{\frac{2h}{g}}. \tag{4.13}$$

Czas lotu t_{lotu} znajdujemy ze wzoru

$$t_{lotu} = \frac{D}{v_0} = \sqrt{\frac{2h}{g}}.$$

Zadanie. Obliczyć prędkość ziarnka grochu wystrzelonego poziomo z procy po obniżeniu się względem pierwotnego poziomu o $h = 1$ m, jeżeli wiadomo, że zasięg w tym momencie wynosi $D = 20$ m. Odp.: ok. 45 m/s.

Wyniki, jakie uzyskaliśmy badając rzut poziomy, wykorzystamy do obliczenia tzw. *pierwszej prędkości kosmicznej*. Tak nazywamy najmniejszą prędkość, z jaką należy wyrzucić jakieś ciało, by mogło ono krążyć dokoła Ziemi. Rys. 4.4 wyjaśnia, co się dzieje, gdy z punktu A wyrzucamy poziomo ciało nadając mu kolejno coraz większe prędkości

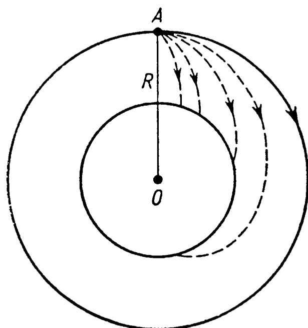
początkowe. Jak widać, przy pewnej granicznej prędkości v_I ciało podczas „spadania" już nie osiąga powierzchni Ziemi, lecz zaczyna krążyć po torze kołowym o promieniu R. Ponieważ R jest niedużo większe od promienia Ziemi (na rysunku skala odległości nie jest zachowana), przeto można uważać, że ciężar krążącego ciała równa się mg. Ciężar ten stanowi siłę dośrodkową, a zatem

$$\frac{mv_I^2}{R} = mg,$$

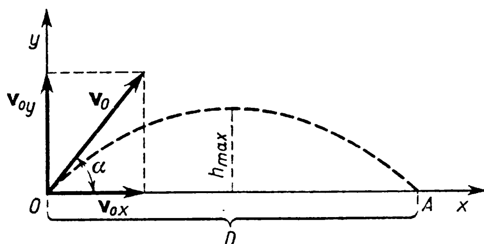
skąd

$$v_I = \sqrt{Rg}. \tag{4.14}$$

Podstawiając jako R przybliżoną wartość promienia Ziemi ($64 \cdot 10^5$ m), otrzymujemy dla pierwszej prędkości kosmicznej zaokrągloną wartość 8 km/s. (W naszych rozważaniach nie braliśmy pod uwagę oporów występujących podczas ruchu.)



Rys. 4.4



Rys. 4.5

4. Rozpatrzmy jeszcze *rzut ukośny*. Nadajemy ciału prędkość początkową $\mathbf{v}_0$ skierowaną pod kątem α względem poziomu (rys. 4.5). Możemy potraktować tę prędkość jako sumę geometryczną prędkości poziomej $\mathbf{v}_{0x}$ o wartości liczbowej

$$v_{0x} = v_0 \cos\alpha$$

i prędkości pionowej $\mathbf{v}_{0y}$ o wartości liczbowej

$$v_{0y} = v_0 \sin\alpha.$$

Ponieważ w kierunku poziomym nie działa na ciało żadna siła, więc składowy ruch poziomy jest prostoliniowy i jednostajny. Prędkość w tym ruchu jest stała i równa v_{0x}. W kierunku pionowym działa siła ciężkości wywołując powstanie ruchu jednostajnie zmiennego, o prędkości v_y zmieniającej się z biegiem czasu. A zatem

$$v_x = v_{0x} = v_0 \cos\alpha,$$

$$v_y = v_{0y} - gt = v_0 \sin\alpha - gt.$$

Drogi x i y przebyte w czasie t odpowiednio w ruchu poziomym i pionowym równają się

$$x = v_0 \cos\alpha \cdot t,$$

$$y = v_0 \sin\alpha \cdot t - \frac{gt^2}{2}.$$

Stąd po wyrugowaniu t otrzymujemy

$$y = x \operatorname{tg}\alpha - \frac{g}{2v_0^2\cos^2\alpha}x^2. \tag{4.15}$$

Jak wiadomo, zależność tego typu daje na wykresie parabolę (rys. 4.5).

W rzucie ukośnym interesującymi wielkościami są: *zasięg rzutu D*, *maksymalna wysokość* h_{max} oraz *czas lotu*. Obliczymy te wielkości i przeprowadzimy dyskusję wyników.

Równanie parabolicznego toru rzutu ukośnego można przepisać w postaci

$$y = -ax^2 + bx, \tag{4.16}$$

gdzie

$$a = \frac{g}{2v_0^2\cos^2\alpha}, \qquad b = \operatorname{tg}\alpha.$$

Obliczenie zasięgu D sprowadza się do obliczenia wartości x dla $y = 0$, czyli

$$(-ax+b)x = 0, \qquad x_1 = 0, \qquad x_2 = \frac{b}{a}.$$

Pierwiastek x_2 jest właśnie szukanym zasięgiem D rzutu:

$$D = x_2 = \frac{\operatorname{tg}\alpha \cdot 2v_0^2\cos^2\alpha}{g},$$

$$D = \frac{v_0^2 \sin 2\alpha}{g}. \tag{4.17}$$

Wobec stałości v_0 i g maksimum zasięgu otrzymuje się przy $\alpha = 45°$. Jednakowe zasięgi odpowiadają (przy pozostałych warunkach niezmiennych) 2 kątom różniącym się od 45° o tę samą liczbę stopni np. 30° i 60°; 40° i 50° itp. (Z jakiej zależności trygonometrycznej to wynika?)

W celu obliczenia maksymalnej wysokości rzutu ukośnego znajdujemy pierwszą pochodną dy/dx i przyrównujemy ją do zera (ogólna metoda wyznaczania maksimum funkcji):

$$\frac{dy}{dx} = -2ax+b, \qquad -2ax+b = 0, \qquad x = \frac{b}{2a}.$$

Jak widać z porównania z poprzednio znalezioną wartością $D = x_2 = b/a$, maksimum funkcji przypada w połowie zasięgu.

Znalezioną wartość x podstawiamy do podstawowego równania (4.16) i znajdujemy

$$y_{max} = h_{max} = -a\frac{b^2}{4a^2} + \frac{b^2}{2a} = \frac{b^2}{4a}.$$

Stąd

$$h_{max} = \frac{\operatorname{tg}^2\alpha \cdot 2v_0^2\cos^2\alpha}{4g},$$

$$h_{max} = \frac{v_0^2\sin^2\alpha}{2g}. \tag{4.18}$$

Analizując h_{max} w zależności od kąta wyrzucenia α stwierdzamy, że maksymalna wysokość odpowiada oczywiście granicznemu przypadkowi rzutu ukośnego pod kątem $\alpha = 90°$ (rzut pionowy).

Całkowity czas trwania lotu można obliczyć np. ze znanej już wartości zasięgu D, przedstawiającej drogę przebytą ruchem jednostajnym z prędkością $v_{0x} = v_0 \cos\alpha$.

$$t_{lotu} = \frac{D}{v_{0x}} = \frac{v_0^2 \sin 2\alpha}{g v_0 \cos\alpha},$$

$$t_{lotu} = \frac{2v_0 \sin\alpha}{g}. \tag{4.19}$$

Równania (4.17)–(4.19) przedstawiają podstawowe wielkości charakteryzujące rzut ukośny.

Wartość liczbową prędkości w dowolnej chwili t znajdujemy z wzoru

$$v = \sqrt{v_x^2 + v_y^2} = \sqrt{v_{0x}^2 + (v_{0y} - gt)^2}.$$

Na podstawie powyższej zależności można sprawdzić, że wartość liczbowa prędkości końcowej v_k jest równa v_0.

$$v_k = \sqrt{v_{0x}^2 + (v_{0y} - gt_{lotu})^2},$$

$$v_k = \sqrt{v_0^2 \cos^2\alpha + \left(v_0 \sin\alpha - g\frac{2v_0 \sin\alpha}{g}\right)^2},$$

$$|v_k| = |v_0|.$$

Otrzymany wynik warto przeanalizować graficznie, aby odpowiedzieć na pytanie, czy wektor prędkości początkowej $\mathbf{v}_0$ jest równy wektorowi prędkości końcowej $\mathbf{v}_k$.

W dotychczasowych rozważaniach badaliśmy wpływ zmian kąta α na zasięg i wysokość rzutu przy niezmiennych v_0 i g. Dla sportowców interesująca jest także inna zależność, a mianowicie: zasięgi rzutów o tej samej prędkości v_0 i tym samym kącie α są różne w różnych miejscowościach na Ziemi wobec zależności D od przyspieszenia ziemskiego g. Łatwo sprawdzić, że różnica procentowa zasięgów rzutów równa się procentowej różnicy wartości g. Ile wynosiłaby ta różnica w przypadku granicznym, gdyby zawody sportowe odbywały się raz na biegunie ($g = 9,83$ m/s^2), a następnie na równiku ($g = 9,78$ m/s^2)?

Omawianie ruchu ciał w polu grawitacyjnym Ziemi zakończymy obliczeniem tzw. *prędkości ucieczki* dowolnego ciała z otoczenia Ziemi. Rzucając ciało pionowo w górę nadajemy mu energię kinetyczną, której kosztem wykonuje ono pracę pokonywania siły ciężkości, co wiąże się ze wzrostem jego energii potencjalnej. Jeśli udzielona energia kinetyczna nie jest dostatecznie duża, to wystarcza jej tylko na osiągnięcie przez ciało skończonej odległości r od środka Ziemi, po czym ciało spada. Aby umożliwić mu oddalenie się aż do nieskończoności, należy mu nadać energię kinetyczną co najmniej równą różnicy $|E_{pR_Z} - E_{p\infty}|$. Ponieważ $E_{p\infty} = 0$, a $E_{pR_Z} = -G\frac{m_Z m}{R_Z}$,

więc

$$E_k \geqslant G\frac{m_Z m}{R_Z}, \quad v \geqslant \sqrt{\frac{2Gm_Z}{R_Z}}.$$

Po podstawieniu znanych danych liczbowych otrzymuje się wartość ok. 11 100 m/s jako minimalną prędkość (zwaną właśnie *prędkością ucieczki* lub *drugą prędkością kosmiczną*), którą należy nadać ciału wyrzuconemu z powierzchni Ziemi, aby (przy zaniedbaniu oporu powietrza) już na Ziemię nie wróciło. Można wykazać, że ciało wyrzucone z prędkością początkową równą drugiej prędkości kosmicznej lub większą od niej oddala się od Ziemi po torze parabolicznym lub hiperbolicznym.

Pytania i zadania

1. Jaka jest treść prawa powszechnego ciążenia?

2. Jaki jest sens fizyczny stałej grawitacyjnej? Jakie są konsekwencje tego, że jej wartość liczbowa jest rzędu 10^{-11} w układzie SI?

3. Podać przykłady stosowalności prawa powszechnego ciążenia do ciał niebieskich.

4. Jaka jest treść praw Keplera? Czy w związku z prawem drugim można by mówić o stałej prędkości polowej? Jak należałoby zdefiniować prędkość polową?

5. Jaki istnieje związek między ciężarem ciał na Ziemi i prawem powszechnego ciążenia.

6. Czy ciężar ciała o masie 1 kg jest stały, czy zależny od położenia względem środka Ziemi? Jeśli zachodzi ta druga ewentualność, to podać ciężar tego ciała w jednostkach SI, gdy ono się znajduje a) na biegunie, b) na równiku, c) w środku Ziemi, d) w odległości równej odległości Księżyca od Ziemi.

7. Jakie dane są potrzebne do obliczenia masy Słońca?

8. Czy zmienność przyspieszenia ziemskiego na Ziemi wiąże się tylko z jej kształtem?

9. Czy wzory obowiązujące przy swobodnym spadku ciał przydają się przy badaniu rzutu poziomego i ukośnego? Dokładnie wyjaśnić.

10. Jakie elementy można wyróżnić przy dokładnej analizie rzutu ukośnego? Wyprowadzić równanie toru oraz podstawowe wzory.

11. Jaka wielkość a) wektorowa, b) skalarna służy do określania pola sił grawitacyjnych? Podać dokładne definicje obu wielkości.

12. Podać definicje I i II prędkości kosmicznej i wyjaśnić sposób ich obliczania.

13. Jak można wyjaśnić prawie całkowity brak wodoru w atmosferze ziemskiej? (Uwzględnić możliwość osiągania dużych prędkości przez lekkie cząsteczki gazowe w czasie zderzeń termicznych w górnych warstwach atmosfery.)

14. Czy prędkość ucieczki ze Słońca jest większa czy też mniejsza niż z Ziemi? Czy przemawia to za obecnością wodoru w atmosferze Słońca?

15. Wyjaśnić, dlaczego całkowita energia planety w jej ruchu dokoła Słońca ma a) stałą, b) ujemną wartość?

16. Wiedząc, że stosunek średnich promieni Ziemi i Marsa w przybliżeniu równa się 2 oraz że masa Marsa równa się $0{,}107m_Z$ (m_Z — masa Ziemi) obliczyć a) przyspieszenie grawitacyjne na Marsie, b) prędkość ucieczki z Marsa.

Odp. $0{,}214g_Z$, ok. 5000 m/s.

17. Wiedząc, że $m_Z/m_K = 81$ obliczyć stosunek odległości od środka Ziemi i środka Księżyca takiego punktu, w którym natężenie pola grawitacyjnego tych dwóch ciał równa się zeru.

Odp. 9:1.

18. Ciała A i B o masach 0,2 kg i 0,8 kg znajdują się w odległości 0,12 m od siebie. Obliczyć pracę potrzebną do przesunięcia masy jednostkowej z punktu leżącego w odległości 0,04 m od masy A do punktu

leżącego w odległości 0,04 m od masy B. Jakie jest w każdym z tych punktów natężenie pola i potencjał grawitacyjny?

Odp. 1) $-5 \cdot 10^{-10}$ J, 2) 0 i ok. $3 \cdot 10^{-8}$ N/kg, 3) -10^{-9} i $1{,}5 \cdot 10^{-9}$ J/kg.

19. Obliczyć stosunek zasięgów i wysokości rzutów ukośnych o tych samych prędkościach początkowych v_0 i przyspieszeniach g, jeśli kąty rzutu wynoszą odpowiednio 30°, 45°, 60°. Przedstawić te trzy rzuty na wspólnym wykresie.

Odp. 1) $\sqrt{3}/2 : 1 : \sqrt{3}/2$, 2) 1:2:3.

20. Z wylotu rury wypływa strumień wody z prędkością v_0 pod kątem α względem poziomu. Na jakiej wysokości h trafi on w pionową ścianę ustawioną w odległości d od wylotu strugi? Jaka jest maksymalna wysokość i jaki byłby zasięg, gdyby usunąć ścianę. Opory zaniedbać.

Odp. $h = d\,\mathrm{tg}\,\alpha - \dfrac{gd^2}{2v_0^2\cos^2\alpha}$.

21. Kulka stalowa spada swobodnie z wysokości h na sprężystą podstawę nachyloną pod kątem α względem poziomu. W jakiej odległości kulka uderzy o podstawę po raz drugi?

Odp. $d = 8h\sin\alpha$.

22. Z jaką prędkością poziomą należy wyrzucić ciało na wysokości h = 1000 km nad Ziemią, aby wytworzył się nowy satelita ziemski? Promień Ziemi przyjąć równy 6400 km.

Odp. ok. 7,4 km/s.

ROZDZIAŁ 5

Dynamika ruchu obrotowego bryły sztywnej

5.1. Rozważania wstępne*

W rozdziale 2 zapoznaliśmy się z prawami dynamiki ruchu punktu materialnego i ruchu postępowego układu punktów materialnych. W odniesieniu do pojedynczego punktu lub do środka masy układu spełnione były zależności

$$m\mathbf{a} = \mathbf{F}, \qquad m_0\mathbf{a} = \sum \mathbf{F}_z.$$

Wynikało z nich jednoznacznie, że za zmiany prędkości v w ruchu postępowym (istnienie przyspieszenia) odpowiedzialna jest siła, zaś miarą bezwładności jest masa m punktu lub masa m_0 układu.

W dynamice ruchu obrotowego bryły sztywnej obowiązują inne zależności: aczkolwiek istnienie siły jest warunkiem koniecznym do wystąpienia zmian prędkości kątowej ω w ruchu obrotowym (istnienie przyspieszenia kątowego α), ale nie jest warunkiem wystarczającym, tzn. nie każda siła wywołuje takie zmiany. Co więcej — zmiany wywołane działaniem siły zależą nie tylko od jej wartości liczbowej, lecz także od punktu przyłożenia i kierunku działania. Wyobraźmy sobie np. koło rowerowe umieszczone poziomo i tak zamocowane, że może się obracać dokoła stałej osi pionowej przechodzącej przez środek. Rozważać będziemy tylko siły działające w płaszczyźnie prostopadłej do osi obrotu. Siła działająca wzdłuż szprychy nie wywoła ruchu obrotowego koła ani też ruchu już istniejącego nie przyspieszy, ani opóźni (będzie tylko działała na łożyska, w jakich osadzona jest oś). Wymienione zmiany wywoła natomiast siła działająca prostopadle do szprychy (a równocześnie stycznie do okręgu zataczanego przez punkt przyłożenia siły), przy czym efekty ilościowe będą różne, zależne od odległości punktu przyłożenia danej siły od osi obrotu (na końcu szprychy, w połowie jej długości, w jednej czwartej itd.). Największe przyspieszenie kątowe odpowiada największemu oddaleniu punktu przyłożenia siły od osi obrotu. Jeśli w dalszych badaniach nie zmieniając ani wartości liczbowej siły, ani punktu jej przyłożenia zmieniać będzie-

---

* Rozważania zawarte w § 5.1–5.10 dotyczą ruchu obrotowego bryły sztywnej dokoła osi obrotu, mającej (dzięki odpowiednim więzom) stałe położenie w przestrzeni.

my kąt, jaki tworzy ona ze szprychą (z promieniem zataczanego okręgu), utrzymując ją stale w płaszczyźnie prostopadłej do osi obrotu, to również zmiany przez nią wywołane będą różne: najsilniejsze — przy ustawieniu prostopadłym do promienia, zerowe — przy ustawieniu wzdłuż promienia.

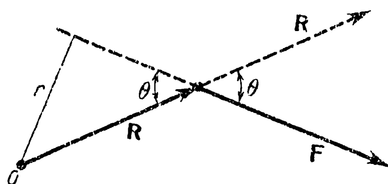
Zależność zmian ω w ruchu obrotowym od wszystkich wymienionych w omawianym przykładzie czynników uwzględnimy przyjmując, że za zmiany w ruchu obrotowym odpowiedzialna jest nie siła, lecz *moment siły względem osi obrotu.* (Definicję momentu siły podamy w następnym punkcie). Jest to jedna z zasadniczych różnic między dynamiką ruchu postępowego i ruchu obrotowego. Jak zobaczymy w § 5.5, miara bezwładności bryły w ruchu obrotowym również jest inna niż w ruchu postępowym. Nie jest nią całkowita masa bryły, lecz tzw. *moment bezwładności*, wielkość uwzględniająca rozmieszczenie mas poszczególnych cząstek bryły względem osi obrotu.

5.2. Moment siły

Moment siły **F** *względem punktu O* definiujemy jako iloczyn wektorowy $\mathbf{R}\times\mathbf{F}$, przy czym wektor **R** jest wektorem łączącym punkt O z punktem przyłożenia siły **F** (rys. 5.1). Moment siły **F** względem punktu O oznaczać będziemy symbolem $\mathbf{M}_O(\mathbf{F})$.

$$\mathbf{M}_O(\mathbf{F}) = \mathbf{R} \times \mathbf{F}. \tag{5.1}$$

Według definicji iloczynu wektorowego wartość liczbowa iloczynu $\mathbf{R}\times\mathbf{F}$ jest równa $RF\sin\theta$, gdzie θ jest kątem między wektorem **R** i wektorem **F** (kąt θ mierzymy wtedy, gdy oba wektory są tak przesunięte, że ich początki przypadają w tym samym punkcie). Z rysunku 5.1 widać, że iloczyn $R\sin\theta = r$, tzn. równa się odległości punktu O od prostej działania siły **F**. Tę odległość nazywamy *ramieniem siły* **F** *względem punktu O*. Możemy więc powiedzieć, że wartość liczbowa momentu siły **F** względem punktu O równa się iloczynowi wartości tej siły przez jej ramię względem punktu O.



Rys. 5.1

Przechodzimy do określenia kierunku i zwrotu momentu siły. Obowiązuje *reguła korkociągu*. Korkociąg ustawiamy prostopadle do płaszczyzny zawierającej **R** i **F** opierając jego ostrze w punkcie O. Następnie rączkę korkociągu obracamy od położenia równoległego do wektora **R** do położenia równoległego do wektora **F** o kąt mniejszy od 180°. Kierunek przesuwania się korkociągu towarzyszący temu obrotowi wyznacza zwrot momentu siły. Przy ustawieniu wektorów **R** i **F** jak na rys. 5.1 moment siły **F** jest skierowany prostopadle za płaszczyznę rysunku.

Łatwo sprawdzić, że ***moment siły względem każdego punktu leżącego na prostej jej działania równa się zeru***, gdyż wtedy $r = 0$.

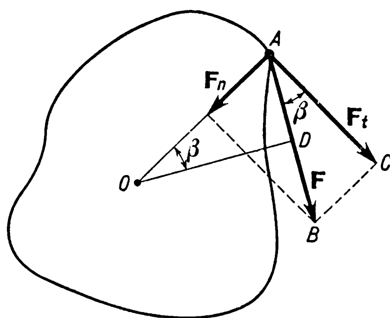
Pojęcie momentu kojarzymy w fizyce nie tylko z siłą, lecz i z innymi wielkościami fizycznymi. Tak np. w § 5.7 będziemy mówili o momencie pędu $\mathbf{R}\times m\mathbf{v}$.

5.3. Warunek powstawania ruchu obrotowego zmiennego

Rozważając w § 1.6 ruch obrotowy bryły sztywnej dokoła stałej osi z punktu widzenia kinematyki stwierdziliśmy, że każdy z punktów bryły, z wyjątkiem punktów leżących na osi obrotu, odbywa ruch po kole o promieniu równym odległości od osi obrotu. Jeśli ruch obrotowy bryły jest jednostajny, to i każdy z punktów bryły odbywa ruch jednostajny kołowy z prędkością liniową $v = \omega r$. W ruchu tym każdy z punktów bryły ma przyspieszenie dośrodkowe i podlega działaniu siły dośrodkowej. Siła ta nie wywołuje zmian wartości liczbowej prędkości.

W ruchu obrotowym niejednostajnym każdy z punktów bryły ma przyspieszenie normalne i przyspieszenie styczne. Odpowiednio do tych dwóch składowych przyspieszenia możemy mówić o sile stycznej do toru i normalnej do toru, skierowanej wzdłuż promienia. Siła normalna wywołuje zakrzywienie toru. Siła styczna jest czynnikiem odpowiedzialnym za zmiany wartości liczbowej prędkości w ruchu obrotowym.

Doszliśmy więc do warunku określającego możliwość powstawania ruchu obrotowego niejednostajnego. Nie wystarcza istnienie siły normalnej, działającej wzdłuż promienia. Na dowolny punkt bryły obracającej się (lub na szereg jej punktów) musi działać ponadto siła zewnętrzna styczna do toru, czyli, innymi słowy, dowolny punkt bryły musi być poddany działaniu takiej siły, która ma składową styczną. Sztywność bryły jest czynnikiem decydującym o występowaniu sił stycznych wewnętrznych, działających na pozostałe punkty bryły.



Rys. 5.2

Rozpatrzmy te zależności z punktu widzenia momentów działających sił. Niech bryła sztywna odbywa ruch obrotowy dokoła osi pionowej, przebijającej płaszczyznę rysunku w punkcie O (rys. 5.2). Na punkt A tej bryły niech działa w płaszczyźnie rysunku siła $\mathbf{F}$ skierowana ukośnie względem promienia wodzącego OA. Tę siłę rozkładamy na składową styczną $\mathbf{F}_t$ i składową normalną $\mathbf{F}_n$. Z punktu O spuszczamy

prostopadłą na kierunek siły **F**, czyli znajdujemy jej ramię OD względem punktu O. Łatwo sprawdzić, że trójkąty AOD i BAC są podobne, a zatem

$$F_t:F = OD:OA,$$

czyli

$$F_t \cdot OA = F \cdot OD. \tag{5.2}$$

Rozpatrzmy oba iloczyny. Odcinek OA jest odległością siły $\mathbf{F}_t$ od osi obrotu, czyli jest ramieniem siły $\mathbf{F}_t$. Podobnie OD jest ramieniem siły **F** względem osi obrotu. Równanie (5.2) stwierdza równość wartości liczbowych momentów sił **F** i $\mathbf{F}_t$ względem osi obrotu. Oba te momenty skierowane są równolegle do osi obrotu za płaszczyznę rysunku.

Moment składowej normalnej $\mathbf{F}_n$ względem osi obrotu równa się zeru, gdyż kierunek tej siły przechodzi przez punkt O.

Streszczając wyniki powiemy:

1. Zmiany w ruchu obrotowym może wywołać tylko taka siła, której moment względem osi obrotu nie równa się zeru.

2. Mimo, że siły **F** i $\mathbf{F}_t$ są różne, to jednak wywołują one jednakowe zmiany w ruchu obrotowym, gdyż tylko składowa styczna $\mathbf{F}_t$ siły **F** powoduje zmiany wartości liczbowej prędkości ruchu obrotowego. Obie te siły mają jednakowe momenty (por. równanie 5.2) względem osi obrotu. Możemy zatem przyjąć, że o *zmianach w ruchu obrotowym decyduje moment siły, a mianowicie jego wartość i kierunek.*

Z tych rozważań widać, że w ruchu obrotowym bryły sztywnej obowiązują inne zależności niż w ruchu postępowym.

Ruch postępowy prostoliniowy jest jednostajny wtedy, gdy na ciało nie działa żadna siła lub wypadkowa sił działających równa się zeru.

Ruch obrotowy jest jednostajny wtedy, gdy wypadkowy moment względem osi obrotu wszystkich sił działających na ciało równa się zeru.

Ruch postępowy prostoliniowy jest zmienny, gdy ciało podlega działaniu siły.

Przy ruchu obrotowym działanie siły jest warunkiem koniecznym, ale niewystarczającym do wywołania *ruchu obrotowego zmiennego*. Tylko taka siła działająca na ciało obracające się wywoła zmianę prędkości kątowej, czyli wystąpienie przyspieszenia kątowego α, której moment względem osi obrotu nie równa się zeru ($\mathbf{M} \neq 0$).

Jeżeli moment siły jest skierowany zgodnie z wektorem $\boldsymbol{\omega}$, to ruch obrotowy jest przyspieszony. Przy przeciwnym zwrocie momentu działającej siły powstaje ruch obrotowy opóźniony.

Poprzednio w rozważaniach kinematycznych ustaliliśmy zależność między składową styczną przyspieszenia a_t i przyspieszeniem kątowym α:

$$a_t = \alpha r.$$

Wiemy też z drugiej zasady dynamiki, że siła $\mathbf{F} = m\mathbf{a}$. Ostatnio dowiedliśmy, że moment siły **F** (rys. 5.2) równa się momentowi jej składowej stycznej $\mathbf{F}_t$. W odniesieniu do pojedynczego punktu materialnego odbywającego ruch po okręgu możemy więc napisać

$$M_O(F) = rF_t = rma_t,$$

czyli

$$M_O(F) = mr^2\alpha. \tag{5.3}$$

5.4. Ruch obrotowy zmienny bryły sztywnej

Od rozważań związanych z działaniem siły **F** na punkt materialny przejdźmy do rozważań ogólniejszych i potraktujmy bryłę sztywną jako zbiór punktów materialnych o masach $m_1, m_2, m_3, \ldots, m_n$, odległych od osi obrotu o $r_1, r_2, r_3, \ldots, r_n$ i podlegających działaniu sił $\mathbf{F}_1, \mathbf{F}_2, \mathbf{F}_3, \ldots, \mathbf{F}_n$. Każda z tych sił wywołuje moment względem osi obrotu $\mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3, \ldots, \mathbf{M}_n$. Momenty wszystkich sił są równoległe do osi obrotu; mogą tylko różnić się zwrotem. Momentom skierowanym zgodnie z wektorem $\boldsymbol{\omega}$ przypisujemy znak plus, skierowanym przeciwnie — znak minus. Bryła jako całość porusza się z określonym przyspieszeniem kątowym α. Dla nadania tego przyspieszenia punktowi materialnemu o masie m_i, umieszczonemu w odległości r_i od osi obrotu, potrzebny jest moment (por. 5.3)

$$M_O(F_i) = m_i r_i^2 \alpha.$$

Moment ten może być wynikiem bezpośredniego działania siły zewnętrznej na i-ty punkt materialny lub też może być wynikiem oddziaływania sąsiednich punktów. Suma geometryczna wszystkich momentów sprowadza się — wobec ich jednakowego kierunku — do sumy algebraicznej. Możemy więc napisać, że moment wypadkowy M_w równa się

$$M_w = m_1 r_1^2 \alpha + m_2 r_2^2 \alpha + \ldots + m_n r_n^2 \alpha$$

lub

$$\mathbf{M}_w = \alpha \sum_{i=1}^{n} m_i r_i^2. \tag{5.4}$$

5.5. Moment bezwładności

Zestawmy zależności obowiązujące przy ruchu postępowym i obrotowym. Siła wypadkowa **F** wywołująca przyspieszenie **a** spełnia równanie $\mathbf{F} = m\mathbf{a}$. Moment wypadkowy siły, wywołujący przyspieszenie kątowe α w ruchu obrotowym, spełnia zależność

$$\mathbf{M}_w = \alpha \sum_{i=1}^{n} m_i r_i^2.$$

Porównując te dwa wzory stwierdzimy, że wyrażenie

$$\sum_{i=1}^{n} m_i r_i^2$$

odgrywa w ruchu obrotowym podobną rolę jak masa w ruchu postępowym. *Suma iloczynów mas poszczególnych cząstek bryły i kwadratów ich odległości od osi obrotu jest*

miarą bezwładności bryły w ruchu obrotowym i nosi nazwę momentu bezwładności J względem danej osi obrotu

$$I = \sum_{i=1}^{n} m_i r_i^2. \tag{5.5}$$

W przypadku ciała sztywnego o ciągłym rozkładzie masy dzielimy je w myśli na nieskończenie małe elementy masy dm i sumowanie we wzorze (5.5) zastępujemy całkowaniem. Wtedy moment bezwładności wyraża się wzorem

$$I = \int r^2 dm, \tag{5.6}$$

przy czym całkowanie rozciągnięte jest na całą objętość ciała.

Ze wzoru (5.5) widać, że o bezwładności bryły obracającej się nie decyduje suma mas poszczególnych cząstek bryły. Zasadnicze znaczenie ma rozmieszczenie mas względem osi obrotu.

Po uwzględnieniu (5.5) wyrażenie (5.4) przyjmuje postać

$$\mathbf{M}_w = \alpha I. \tag{5.7}$$

Jest to podstawowe prawo ruchu obrotowego. Zgodnie z tym prawem *wypadkowy moment siły działający na ciało obracające się równa się iloczynowi momentu bezwładności względem aktualnej osi obrotu przez przyspieszenie kątowe obrotu bryły.* (Prawo w tej prostej postaci obowiązuje, gdy moment bezwładności ciała jest niezmienny w czasie.)

Obliczanie momentów bezwładności brył geometrycznych względem różnych osi jest zagadnieniem rozwiązywanym przez całkowanie. Podamy niżej kilka wyników takich obliczeń:

Kula jednorodna o promieniu R i masie m; oś przechodzi przez środek kuli: $I = \frac{2}{5} mR^2$.

Równoległościan prostokątny o masie m i bokach a, b, c; oś równoległa do wysokości, przechodząca przez środek podstawy: $I = \frac{1}{12}(a^2+b^2)m$.

Walec prosty kołowy o masie m, wysokości h i promieniu R; podłużna oś symetrii: $I = \frac{1}{2} mR^2$.

Taki sam walec; poprzeczna oś symetrii: $I = \left(\frac{R^2}{4} + \frac{h^2}{12}\right)m$.

Walec wydrążony o masie m, promieniu zewnętrznym R i wewnętrznym r; podłużna oś symetrii: $I = \frac{1}{2}(r^2+R^2)m$.

Jak widać, w każdym z wyrażeń określających moment bezwładności bryły występuje całkowita masa bryły m pomnożona przez jakiś czynnik. Można zatem moment bezwładności przedstawić jako

$$I = mk^2.$$

Czynnik k nosi nazwę *ramienia bezwładności*. Jest to wielkość, która po podniesieniu do kwadratu i pomnożeniu przez całkowitą masę bryły daje moment bezwładności.

Dla podkreślenia zależności momentu bezwładności od wyboru osi podajemy twierdzenie Steinera*. Zgodnie z tym twierdzeniem, moment bezwładności I względem do-

* Dowód tego twierdzenia można znaleźć w każdym obszerniejszym podręczniku mechaniki.

wolnej osi jest związany z momentem bezwładności I_0 względem osi przechodzącej przez środek masy i równoległej do osi danej następującą zależnością:

$$I = I_0 + md^2,$$

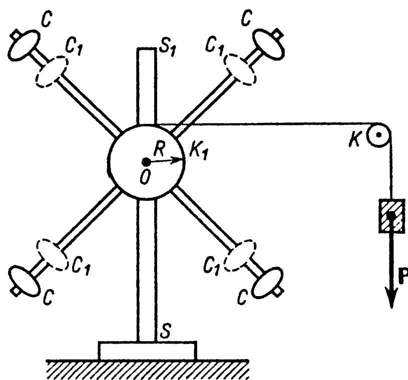
gdzie m oznacza całkowitą masę bryły, d — odległość wzajemną obu osi.

Rozpatrzmy następujące doświadczenie. Na pionowym statywie SS_1 (rys. 5.3) osadzony jest czteroramienny krzyżak w ten sposób, że może być wprawiony w ruch obrotowy dokoła osi poziomej O. Na każdym z ramion krzyżaka umieszczony jest przesuwalny ciężarek C. Do wywołania ruchu obrotowego służy odważnik o ciężarze P, zawieszony na nitce przerzuconej przez krążek K. Siła **P** działając stycznie do obwodu drugiego krążka K_1 o promieniu R wywołuje stały moment siły równy liczbowo RP względem osi obrotu. Ten moment siły $M(P)$ powoduje ruch obrotowy przyspieszony zgodnie z równaniem

$$M(P) = I\alpha.$$

Przy danym $M(P)$ wartość przyspieszenia kątowego zależy od rozmieszczenia ciężarków C. Gdy są one umieszczone dalej od osi obrotu, moment bezwładności bryły jest większy, a tym samym przyspieszenie kątowe jest mniejsze. Przesunięcie ciężarków do położenia C_1 (zaznaczonego linią przerywaną na rys. 5.3) powoduje zmniejszenie momentu bezwładności, a co za tym idzie wzrost α przy danym, stałym momencie $M(P)$.

Jeśli kolejno przy obu położeniach ciężarków C siła będzie działała na bryłę równie długo, to wywoła różne końcowe prędkości kątowe: większą — w przypadku mniejszego momentu bezwładności i odwrotnie.



Rys. 5.3

Prędkość kątowa ω ruchu obrotowego jednostajnie zmiennego bryły zmienia się według wzoru

$$\omega = \omega_0 + \alpha t$$

(gdzie ω_0 oznacza początkową prędkość kątową), analogicznie do wzoru na prędkość w ruchu postępowym jednostajnie zmiennym $v = v_0 + at$. W rozważanym przez nas przypadku $\omega_0 = 0$.

Można też postawić następujące pytanie. Jaki powinien być stosunek czasów t_1 i t_2 działania stałego momentu siły $M(P)$ na dwie bryły o momentach bezwładności I_1 i I_2, aby rozpoczynając ruch od prędkości kątowej $\omega_0 = 0$ osiągnęły one tę samą prędkość kątową ω?

Obowiązują równania

$$\omega = \alpha_1 t_1, \qquad \omega = \alpha_2 t_2.$$

Stąd

$$\frac{t_1}{t_2} = \frac{\alpha_2}{\alpha_1}.$$

Ale

$$\alpha_1 = \frac{M(P)}{I_1} \quad \text{i} \quad \alpha_2 = \frac{M(P)}{I_2},$$

a więc ostatecznie

$$\frac{t_1}{t_2} = \frac{I_1}{I_2}.$$

Czasy działania stałego momentu siły na dwie bryły o momentach bezwładności I_1 i I_2 potrzebne do uzyskania tej samej prędkości kątowej ω są wprost proporcjonalne do momentów bezwładności.

5.6. Energia kinetyczna w ruchu obrotowym

Wychodzimy z ogólnego wzoru na energię kinetyczną punktu materialnego:

$$E_k = \frac{mv^2}{2}.$$

Bryłę sztywną rozpatrujemy znowu jako zbiór dużej liczby punktów materialnych o masach $m_1, m_2, \ldots, m_n$, umieszczonych w odległościach $r_1, r_2, \ldots, r_n$ od osi obrotu. Gdy bryła obraca się z prędkością kątową ω, poszczególne punkty materialne mają odpowiednio prędkości liniowe

$$v_1 = \omega r_1, \qquad v_2 = \omega r_2, \qquad \ldots, \qquad v_n = \omega r_n.$$

Każdy z nich ma więc określoną energię kinetyczną

$$\frac{m_1\omega^2 r_1^2}{2}, \qquad \frac{m_2\omega^2 r_2^2}{2}, \qquad \ldots, \qquad \frac{m_n\omega^2 r_n^2}{2}.$$

Energia kinetyczna całej bryły jest sumą energii kinetycznych wszystkich punktów materialnych wchodzących w skład bryły, a więc

$$E_k = \frac{\omega^2}{2}(m_1 r_1^2 + m_2 r_2^2 + \ldots + m_n r_n^2).$$

Uwzględniając równanie (5.5) znajdujemy

$$E_k = \frac{I\omega^2}{2}. \tag{5.8}$$

A zatem *energia kinetyczna bryły obracającej się dokoła nieruchomej osi równa się połowie*

iloczynu momentu bezwładności i kwadratu prędkości kątowej. Porównując wyrażenie (5.8) ze wzorem $E_k = mv^2/2$ raz jeszcze stwierdzamy, że moment bezwładności odgrywa w ruchu obrotowym podobną rolę jak masa w ruchu postępowym.

Przykład. Obliczyć energię kinetyczną kuli o masie m i promieniu r, toczącej się bez poślizgu z prędkością v. Zagadnienie jest nieco skomplikowane ze względu na to, że tocząca się kula odbywa dwa ruchy: ruch postępowy z prędkością v i ruch obrotowy względem osi przechodzącej przez środek kuli i nie zmieniającej swego położenia w kuli. Energie kinetyczne związane z wymienionymi ruchami wynoszą odpowiednio

$$E_k' = \frac{mv^2}{2}, \quad E_k'' = \frac{I\omega^2}{2} = \frac{2}{5}\,\frac{mr^2\omega^2}{2}.$$

Prędkość kątową ω obrotu kuli chcemy powiązać z prędkością v ruchu postępowego. W tym celu bierzemy pod uwagę punkt styczności kuli z płaszczyzną, po której się ona toczy. Prędkość chwilowa tego punktu równa się zeru, ale równocześnie może być ona przedstawiona jako suma geometryczna prędkości $\mathbf{v}$ ruchu postępowego (skierowanej np. w prawo) i prędkości liniowej $\mathbf{v}'$ tego punktu w jego ruchu obrotowym (skierowanej w lewo). A zatem

$$0 = v - v'.$$

Ale ze wzoru (1.23) wynika, że $v' = \omega r$, a więc

$$v - \omega r = 0, \quad \text{czyli} \quad \omega = \frac{v}{r}.$$

Szukaną energię kinetyczną kuli toczącej się bez poślizgu po płaszczyźnie poziomej można więc przedstawić w postaci

$$E_k = E_k' + E_k'' = \frac{mv^2}{2} + \frac{2mv^2}{5\cdot 2},$$

$$E_k = \frac{7}{5}\,\frac{mv^2}{2}.$$

5.7. Moment pędu (kręt) bryły

Przez *moment pędu punktu materialnego* o masie m poruszającego się z prędkością v względem osi obrotu odległej o r rozumiemy

$$\mathbf{r} \times m\mathbf{v} \tag{5.9}$$

(w tym wzorze wektor $\mathbf{r}$ oznacza wektor promienia wodzącego). W ruchu obrotowym kierunki wektorów $\mathbf{v}$ i $\mathbf{r}$ są wzajemnie prostopadłe, a więc wartość liczbowa momentu pędu równa się rmv.

Zwrot momentu pędu jest, jak łatwo sprawdzić stosując regułę korkociągu, zgodny z wektorem prędkości kątowej (wzdłuż osi obrotu).

Całkowity moment pędu L (czyli tzw. kręt) bryły obracającej się znajdujemy sumując arytmetycznie wszystkie momenty pędu poszczególnych cząstek bryły:

$$L = r_1 m_1 v_1 + r_2 m_2 v_2 + r_3 m_3 v_3 + \ldots + r_n m_n v_n,$$

$$L = \omega(m_1 r_1^2 + m_2 r_2^2 + \ldots + m_n r_n^2),$$

$$L = \omega \sum_{i=1}^{n} m_i r_i^2. \tag{5.10}$$

Uwzględniając to, że mamy do czynienia z wielkościami wektorowymi, napiszemy

$$\mathbf{L} = I\boldsymbol{\omega}. \tag{5:10'}$$

Kręt bryły obracającej się dokoła nieruchomej osi równa się iloczynowi prędkości kątowej i momentu bezwładności bryły względem tej osi.

5.8. Zależność krętu bryły od momentu siły. Zasada zachowania krętu

Zacznijmy znowu od rozważań związanych z pojedynczym punktem bryły. Moment siły F_1, działającej na punkt materialny o masie m_1, odległy od osi obrotu o r_1, wyraża się wzorem

$$M(F_1) = m_1 r_1^2 \alpha.$$

Ale

$$\alpha = \frac{d\omega}{dt},$$

a zatem

$$M(F_1) = m_1 r_1^2 \frac{d\omega}{dt}.$$

Wprowadzając r_1 jako stałą pod znak pochodnej otrzymamy

$$\frac{d(r_1\omega)}{dt} = \frac{dv_1}{dt}.$$

Wprowadzając następnie m_1 i r_1 pod znak pochodnej znajdujemy

$$\frac{d(r_1 m_1 v_1)}{dt}.$$

Iloczyn $r_1 m_1 v_1$ przedstawia wartość liczbową momentu pędu danego punktu bryły (sinus kąta między r_1 i v_1 równa się jedności). A zatem

$$M(F_1) = \frac{d(r_1 m_1 v_1)}{dt},$$

czyli *moment siły równa się pochodnej momentu pędu względem czasu*. Uwzględniając wektorowy charakter wielkości występujących w otrzymanym równaniu można też stwierdzić, że wektor momentu siły i wektor pochodnej momentu pędu względem czasu mają kierunki jednakowe wzdłuż osi obrotu bryły.

Przechodząc od pojedynczego punktu materialnego do całego ich zbiorowiska, stanowiącego bryłę, otrzymamy po lewej stronie równania wypadkowy moment siły, a po stronie prawej pochodną względem czasu krętu bryły:

$$M_w = \frac{d\sum_{i=1}^{n} r_i m_i v_i}{dt}.$$

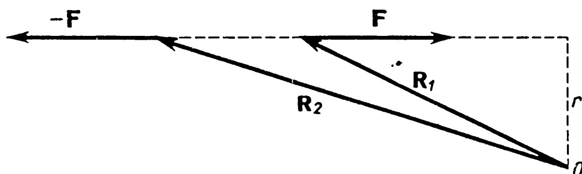
Uwzględniając wektorowy charakter wielkości oraz wprowadzony dla krętu symbol **L** przepiszemy tę zależność w postaci

$$\mathbf{M}_w = \frac{d\mathbf{L}}{dt}. \tag{5.11}$$

Z równania (5.11) wyciągamy wniosek, że *gdy wypadkowy moment siły* $\mathbf{M}_w$ *równa się zeru, to kręt bryły pozostaje stały*. Stanowi to treść tzw. *zasady zachowania krętu*, zwanej też *zasadą zachowania momentu pędu* (odpowiednik zasady zachowania pędu obowiązującej w ruchu postępowym). Kręt bryły może ulec zmianie jedynie pod działaniem momentu siły.

5.9. Wnioski z zasady zachowania krętu

W każdym nieodosobnionym układzie ciał możemy wyróżnić siły wewnętrzne i siły zewnętrzne. Siły wewnętrzne działające między poszczególnymi częściami układu występują zawsze parami, zgodnie z III zasadą dynamiki. Jako równe i przeciwnie skierowane dają one w układzie wypadkową siłę równą zeru. Wykażemy, że ich moment wypadkowy względem dowolnego punktu też równa się zeru (rys. 5.4). Momenty sił **F** i —**F** względem punktu *O* równają się Fr i $-Fr$. Są więc sobie równe liczbowo, mają kierunki jednakowe, lecz zwroty przeciwne: moment siły **F** skierowany jest za płaszczyznę rysunku, moment siły —**F** przed płaszczyznę rysunku. *Siły wewnętrzne nie mają wypadkowego momentu siły, czyli nie mogą powodować zmiany krętu.*



Rys. 5.4

Czy to jednak dowodzi, że w układzie odosobnionym (a więc w takim, w którym nie działają siły zewnętrzne) nie mogą zajść żadne zmiany w ruchu obrotowym ciała? Stałości wypadkowego momentu pędu **L** odpowiada stałość iloczynu $I\omega$ (wzór 5.10′). Jeśli mamy możność zmieniania momentu bezwładności układu pod działaniem sił wewnętrznych, to zmianom I muszą towarzyszyć takie zmiany ω, aby iloczyn $I\omega$ był stały.

Jeśli więc łyżwiarz na lodzie wykonuje piruet, to rozsuwając szeroko ręce zwiększa swój moment bezwładności, a tym samym zmniejsza prędkość kątową obrotu. I odwrotnie — „skupiając" możliwie najbardziej całą swą masę dokoła osi obrotu zmniejsza swój moment bezwładności, co powoduje wzrost ω.

Analogiczne spostrzeżenia poczynimy, obserwując człowieka siedzącego na stołku kręcącym się dokoła osi pionowej i trzymającego w rękach ciężkie kule. Składanie rąk

na piersi, a następnie rozkładanie ich możliwie szeroko powoduje zmianę momentu bezwładności, a równocześnie zmianę prędkości kątowej układu.

Za pomocą takiego samego kręconego stołka można zademonstrować inny jeszcze efekt. Niech człowiek stojący na nieruchomym stołku trzyma w ręku oś pionową, na której u góry umocowane jest koło rowerowe. Jeżeli człowiek wprawi w ruch obrotowy trzymane nad sobą koło, to stołek zacznie się przy tym obracać w stronę przeciwną. Wynik jest zgodny z zasadą zachowania momentu pędu. W rozważanym układzie odosobnionym kręt, równy początkowo zeru, musi zachować tę wartość. A zatem kręt koła $I_1\omega_1$ musi być zrównoważony przeciwnym co do znaku, lecz równym co do wartości krętem $I_2\omega_2$ stołka i stojącego na nim człowieka.

Podobnie przemieszczanie się mas powietrza w pobliżu powierzchni Ziemi związane z różnicami ciśnień powoduje zmiany I układu Ziemia-atmosfera, co pociąga za sobą małe, ale mierzalne zmiany ω. Zmniejszanie się rozmiarów Ziemi podczas stygnięcia jej skorupy również miało wpływ na wartość ω Ziemi.

5.10. Zestawienie podstawowych wielkości i wzorów obowiązujących w ruchu postępowym i obrotowym brył sztywnych

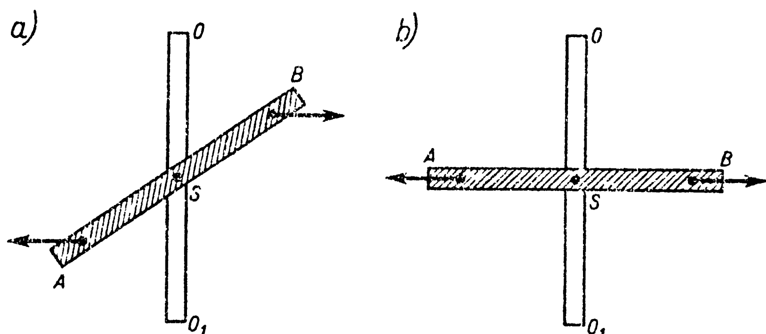
Ruch postępowy	Ruch obrotowy
$v = \frac{ds}{dt}$	$\omega = \frac{d\varphi}{dt}$
$a = \frac{dv}{dt}$	$\alpha = \frac{d\omega}{dt}$
m	I
$\mathbf{F}$	$\mathbf{M}(F)$
$\mathbf{F} = m\mathbf{a}$	$\mathbf{M}(F) = I\boldsymbol{\alpha}$
$\mathbf{p} = m\mathbf{v}$	$\mathbf{L} = I\boldsymbol{\omega}$
$\mathbf{F} = \frac{d(m\mathbf{v})}{dt}$	$\mathbf{M}(F) = \frac{d\mathbf{L}}{dt}$
$E_k = \frac{mv^2}{2}$	$E_k = \frac{I\omega^2}{2}$

5.11. Swobodne osie obrotu

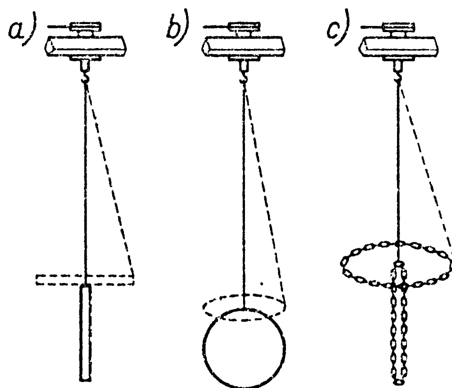
W dotychczasowych rozważaniach nie poświęcaliśmy większej uwagi osiom obrotu zakładając jedynie, że oś — dzięki np. odpowiednim więzom — ma stałe położenie w przestrzeni.

Doświadczenie i teoria wykazują, że w odniesieniu do pewnych osi nie ma istotnego znaczenia ich zamocowanie: ruch obrotowy rozpoczęty przy osi zamocowanej utrzymuje się bez zmian (jeśli można zaniedbać tarcie) po zwolnieniu osi z więzów. Takie osie

obrotu noszą nazwę swobodnych. Badania doświadczalne i teoretyczne wskazują, że *osiami swobodnymi* mogą być osie przechodzące przez środek ciężkości ciała i tak skierowane, by moment bezwładności miał względem nich wartość ekstremalną, tzn. wartość minimalną lub maksymalną. Jeżeli ruch obrotowy odbywa się dokoła osi zamocowanej, przechodzącej przez środek ciężkości, lecz nie będącej równocześnie jedną z możliwych osi swobodnych ciała, to po usunięciu więzów ciało zmienia swe położenie w przestrzeni tak długo, aż osią obrotu stanie się oś swobodna. Czynnikiem wywołującym zmianę położenia w przestrzeni takiej osi po zwolnieniu jej z łożysk jest para sił wytworzona przez siły odśrodkowe (bliższe wiadomości o parze sił — § 6.2). Istnienie takiej



Rys. 5.5



Rys. 5.6

pary sił widać na rys. 5.5a, przedstawiającym obrót pręta AB dokoła osi OO_1, przechodzącej przez środek ciężkości pręta S, lecz nie będącej jego osią symetrii. Gdy oś OO_1 jest zamocowana, zaznaczona na rysunku para sił usiłuje wyrwać ją z łożysk. Gdy oś zostaje zwolniona, to ustawia się ona w stosunku do pręta tak, jak na rys. 5.5b. Staje się ona wtedy swobodną osią obrotu (będąc równocześnie osią symetrii ciała), nie podlegającą już parze sił. Na ogół przy tej zmianie położenia osi zmienia się moment bezwładności ciała, czemu towarzyszy zmiana prędkości kątowej obrotu.

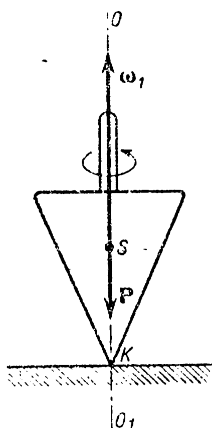
Rysunki 5.6a, b, c przedstawiają wyniki doświadczeń, gdy ciała o różnej postaci (pręt, krążek, łańcuszek) zawieszone na mocnych nitkach, przymocowanych do osi szybkiego motorka, przyjmują takie położenia w przestrzeni, że ich obrót odbywa się dokoła osi swobodnych.

Warto raz jeszcze podkreślić, że swobodną trwałą osią obrotu ciała nie może być dowolna jego oś symetrii. Oś symetrii ciała będzie tylko wtedy osią swobodną, gdy obliczony względem niej moment bezwładności będzie miał wartość ekstremalną.

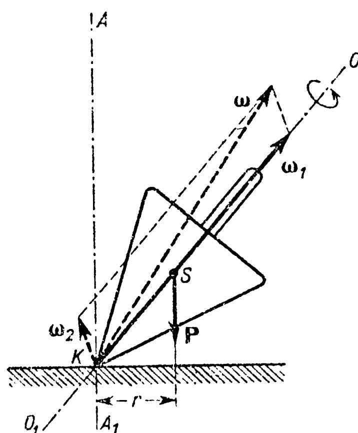
5.12. Ruch precesyjny

W pierwszych punktach tego rozdziału rozważaliśmy przypadek, gdy ciało sztywne odbywa jeden ruch obrotowy dokoła stałej osi. Obecnie przejdziemy do omówienia przypadku składania ruchów obrotowych, w którego wyniku powstaje tzw. *ruch precesyjny*.

Wyobraźmy sobie bąka (rys. 5.7) wirującego dokoła osi symetrii OO_1 ustawionej pionowo. Przy takim ustawieniu bąka siła ciężkości P tylko przyciska go do podstawy, na której jest wsparty. Siłę ciężkości odkładamy od środka ciężkości S pionowo w dół. Kierunek siły ciężkości pokrywa się z kierunkiem osi obrotu, a zatem moment siły ciężkości względem punktu podparcia K równa się zeru. Bąk odbywa tylko jeden ruch obrotowy dokoła stałej osi obrotu OO_1.



Rys. 5.7



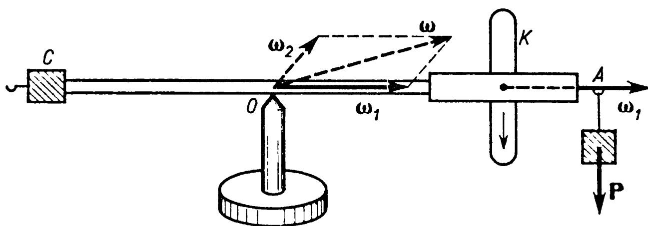
Rys. 5.8

Weźmy pod uwagę inne położenie bąka, takie w którym bąk jest nieco odchylony od pionu (rys. 5.8). Niech bąk odbywa w dalszym ciągu ruch obrotowy dokoła swej osi symetrii OO_1. Tym razem siła ciężkości **P**, przyłożona w punkcie S, działająca pionowo w dół wywołuje moment względem punktu podparcia K równy liczbowo Pr. Moment ten usiłuje wprawić bąka w ruch obrotowy dokoła osi poziomej, prostopadłej do płaszczyzny rysunku, przechodzącej przez punkt K. Możemy więc uważać, że bąk

odbywa dwa ruchy obrotowe: jeden z prędkością kątową ω_1 dokoła osi OO_1, drugi — z prędkością kątową ω_2 dokoła osi poziomej przechodzącej przez punkt K i prostopadłej do płaszczyzny rysunku. Obie prędkości kątowe odkładamy jako wektory wzdłuż osi obrotów, skierowując je odpowiednio do kierunków ruchu obrotowego. Z równoczesności odbywania obu ruchów obrotowych wynika obrót dokoła chwilowej osi obrotu. Kierunek tej „wypadkowej" osi obrotu znajdujemy dodając do siebie geometrycznie wektory ω_1 i ω_2. Otrzymujemy wektor wypadkowy ω. Oś bąka tak się ustawia, aby mieć położenie zgodne z ω. Kierunek wektora ω nie wytycza jednak nowej stałej osi obrotu, gdyż przy tym nowym ustawieniu znowu moment siły ciężkości wywołuje ruch obrotowy dokoła osi poziomej, skierowanej już ukośnie względem płaszczyzny rysunku. Należy więc ponownie znaleźć wypadkową prędkość kątową i tym samym określić nowe położenie chwilowej osi obrotu. W wyniku takiego ciągłego sumowania ruchów obrotowych powstaje ruch tzw. precesyjny.

Bąk wiruje w ten sposób, że obraca się dokoła swej osi symetrii, a równocześnie jego oś symetrii zakreśla stożek dokoła osi pionowej AA_1. W przypadku do którego odnosi się rys. 5.8, oś bąka przesuwa się za płaszczyznę rysunku.

Rozważając działanie siły ciężkości na bąka, trzeba dokładnie zdawać sobie sprawę z odmiennych skutków jej działania w przypadku, gdy bąk jest nieruchomy i w przypadku, gdy wykonuje on ruch obrotowy. Bąk nie wirujący odchylony od położenia pionowego i puszczony swobodnie upada na podstawę pod działaniem siły ciężkości. Bąk wirujący, wytrącony z położenia pionowego, nie upada, lecz zaczyna wykonywać ruch precesyjny. Precesję wywołuje i podtrzymuje moment siły ciężkości względem punktu K, stały co do wartości (przy danym nachyleniu bąka), lecz zmienny co do kierunku. Zmiana kierunku obrotu bąka, czyli zmiana zwrotu ω_1 na przeciwny, wywołuje zmianę kierunku ruchu precesyjnego.



Rys. 5.9

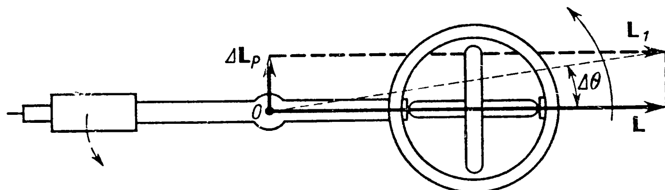
Rozważmy inny przykład powstawania ruchu precesyjnego. Ogólnie znany jest przyrząd zwany wagą żyroskopową (rys. 5.9). Na ostrzu O wspiera się belka. Na jednym jej końcu umieszczony jest ciężarek C, na drugim znajduje się kołowa ramka, w której osadzony jest krążek K. Krążek K może być wprawiony w bardzo szybki ruch obrotowy dokoła osi pokrywającej się z belką. Masa ciężarka C jest tak dobrana, że belka zajmuje położenie poziome. Gdybyśmy na haczyku A zawiesili dodatkowy ciężarek P (przy nieruchomym krążku K), to ramię OA belki obniżyłoby się rozpoczynając ruch obrotowy dokoła osi poziomej, przechodzącej przez punkt O i prostopadłej do płaszczyzny ry-

sunku. Prędkość kątowa ω_2 tego obrotu skierowana byłaby wzdłuż tej osi za płaszczyznę rysunku.

Inaczej się sprawa przedstawia, jeżeli równocześnie krążek K odbywa szybki ruch obrotowy. Przy takim kierunku obrotu, jak zaznaczony na rysunku, wektor prędkości kątowej ω_1 tego ruchu skierowany jest poziomo wzdłuż belki od O w stronę K. Oba wektory ω_1 i ω_2 sumują się geometrycznie dając wypadkową prędkość obrotu ω oraz wytyczając położenie nowej, chwilowej osi obrotu krążka, czyli, innymi słowy, nowe położenie belki.

Ponieważ moment ciężaru **P** działa stale, więc podobnie jak w przypadku bąka ciągle sumują się wektory prędkości kątowych obu składowych ruchów obrotowych. W wyniku otrzymujemy ruch precesyjny wagi żyroskopowej, w czasie którego beleczka CA krąży dokoła pionowej osi przechodzącej przez punkt O. W warunkach odpowiadających rys. 5.9 punkt A beleczki porusza się za płaszczyznę rysunku.

Przez zawieszenie ciężarka P na haczyku przy C zmieniony zostaje kierunek momentu jego siły ciężkości względem punktu O. To prowadzi do zmiany zwrotu ω_2 i do zmiany kierunku ruchu precesyjnego. Zmianę kierunku ruchu precesyjnego można by również wywołać przez zmianę kierunku obrotu krążka K. Równoczesna zmiana kierunku obrotu krążka i przewieszenie ciężarka P na przeciwną stronę nie zmienia kierunku ruchu precesyjnego. Kierunek wypadkowej prędkości ω, określającej położenie chwilowej osi obrotu, pozostaje niezmieniony.



Rys. 5.10

Opierając się na przykładzie wagi żyroskopowej można łatwo znaleźć zależność między prędkością kątową ruchu precesyjnego a momentem siły, wywołującej precesję, oraz krętem obracającego się krążka. Nawiązując do rys. 5.9, na nowym rys. 5.10 (przedstawiającym widok wagi żyroskopowej z góry) zaznaczono wektor krętu **L** obracającego się krążka, skierowany wzdłuż ramienia beleczki od punktu O. Wektor ten ma wartość liczbową $L = I\omega$. Zawieszony w punkcie A (rys. 5.9) ciężarek P wywołuje moment siły spełniający równanie

$$M(P) = \frac{dL_P}{dt},$$

które dla krótkiego czasu Δt można przepisać w postaci

$$\Delta L_P = M(P)\Delta t.$$

Wektor $\Delta \mathbf{L}_P$ przedstawia kręt wywołany działaniem momentu siły $\mathbf{M}(P)$ w czasie Δt. Jest on skierowany prostopadle do beleczki, a więc prostopadle do wektora pierwotnego

krętu **L** krążka (rys. 5.10). Nowo wytworzony kręt $\Delta \mathbf{L}_P$ dodaje się do krętu pierwotnego **L**, dzięki czemu powstaje wypadkowy kręt $\mathbf{L}_1$, leżący również w płaszczyźnie poziomej, obrócony względem pierwotnego kierunku **L** o mały kąt $\Delta\theta$.

Opierając się na przybliżonej równości $\Delta\theta \cong \operatorname{tg}\Delta\theta$, znajdujemy

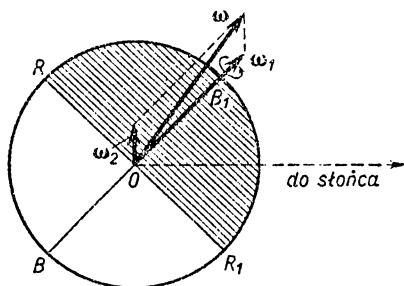
$$\Delta\theta = \frac{\Delta L_P}{L} = \frac{M(P)\Delta t}{I\omega},$$

a zatem prędkość kątowa precesji wynosi

$$\omega_{prec} = \frac{\Delta\theta}{\Delta t} = \frac{M(P)}{I\omega}.$$

Innymi słowy, *prędkość kątowa ruchu precesyjnego jest proporcjonalna do momentu siły wywołującej precesję i odwrotnie proporcjonalna do krętu obracającego się układu.*

Ziemia również wykonuje ruch precesyjny. Obracając się dokoła własnej osi ma ona prędkość kątową ω_1. Oś ziemska nachylona jest względem płaszczyzny ekliptyki pod kątem 66,6°. To powoduje, że w pewnych okresach roku półkula północna znajduje się bliżej Słońca niż południowa (rys. 5.11). Wobec tego Słońce różnie przyciąga obie półkule, w związku z czym występuje moment obrotowy, wywołujący ruch z prędkością kątową ω_2 dokoła osi O prostopadłej do płaszczyzny rysunku. Suma geometryczna wektorów ω_1 i ω_2 określa położenie chwilowej osi obrotu. Powstaje ruch precesyjny, w wyniku którego przedłużenie osi ziemskiej przebija sklepienie niebieskie w punktach układających się na kole zakreślanym jeden raz w ciągu 26 000 lat.

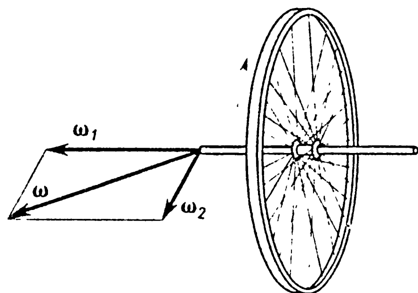


Rys. 5.11

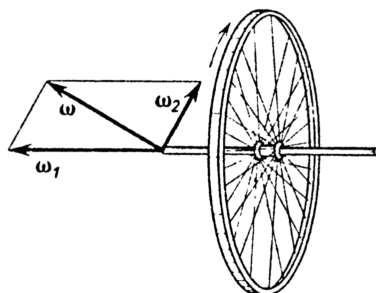
Nieliczni zapewne rowerzyści i motocykliści zdają sobie sprawę z tego, że często wykorzystują ruch precesyjny. Dzieje się to wtedy, gdy przerzucając lekko tułów na lewo lub prawo wywołują odpowiedni skręt roweru czy motocykla.

Rozważmy zjawiska, jakie występują, gdy rower zostaje przechylony. Ruch postępowy koła w tej chwili nas nie interesuje. Obracające się koło, ustawione pionowo, ma prędkość kątową ω_1 skierowaną poziomo. Przechylenie koła prowadzi do powstania momentu obrotowego siły ciężkości względem punktu oparcia koła o nawierzchnię drogi. Występuje więc nowy ruch obrotowy o prędkości kątowej ω_2. Przeanalizujmy te prędkości za pomocą rys. 5.12. Rysunek przedstawia widok koła jadącego w kierunku za-

znaczonym strzałką boczną. Wobec nie zaznaczonego na rysunku pochylenia koła w lewo występuje prędkość kątowa ω_2 o kierunku przeciwnym do ruchu roweru. Sumowanie geometryczne prędkości ω_1 i ω_2 prowadzi do wektora wypadkowego ω określającego kierunek chwilowej osi obrotu. Prostopadle do tego kierunku ustawia się płaszczyzna koła. Koło skręca zatem w lewo, zmieniając kierunek swego ruchu postępowego. Innymi słowy, rowerzysta pochylając się w lewo powoduje skręt koła w lewo. Oś koła pozostaje nadal pozioma. Łatwo się przekonać, że pochylenie koła w prawo zmienia zwrot ω_2 na przeciwny (rys. 5.13) i ostatecznie koło skręca w prawo.



Rys. 5.12



Rys. 5.13

Pytania i zadania

1. Jaki ruch bryły sztywnej nazywamy a) postępowym, b) obrotowym?

2. Czy o zmianach prędkości liniowej w ruchu postępowym i prędkości kątowej w ruchu obrotowym decydują te same czynniki? Omówić je dokładnie.

3. Podać definicję momentu siły.

4. Posługując się pojęciem momentu siły podać warunek powstawania ruchu obrotowego bryły sztywnej a) jednostajnego, b) zmiennego.

5. Co jest miarą bezwładności bryły w ruchu a) postępowym, b) obrotowym? Na czym polega istota różnicy?

6. Jaki wzór w ruchu obrotowym jest odpowiednikiem wzoru $\mathbf{F} = m\mathbf{a}$ obowiązującego w ruchu postępowym? Podać wyprowadzenie tego wzoru.

7. W jaki sposób w doświadczeniu przedstawionym na rys. 5.3 można zmieniać przyspieszenie kątowe bryły?

8. Od jakich czynników zależy energia kinetyczna bryły w ruchu obrotowym? Czy istnieje analogia między wzorami na energie kinetyczne w ruchu postępowym i obrotowym bryły?

9. Jaki jest w ruchu obrotowym odpowiednik wektora pędu występującego w ruchu postępowym?

10. Jaki wzór wiąże wypadkowy moment siły z krętem bryły? Podać wyprowadzenie tego wzoru.

11. Sformułować zasadę zachowania momentu pędu i zilustrować ją przykładami.

12. Zestawić podstawowe wielkości i wzory obowiązujące w ruchu postępowym i obrotowym bryły.

13. Czy przy wyprowadzaniu wzorów na moment siły, kręt i energię kinetyczną w ruchu obrotowym stosuje się podobną metodę? Na czym polega analogia w postępowaniu?

14. Co charakteryzuje swobodne osie obrotu? Omówić doświadczenia wykazujące ich istnienie.

15. Scharakteryzować ruch precesyjny na przykładzie bąka i wagi żyroskopowej.

16. Przeanalizować ruch koła rowerowego na zakręcie.

17. Przez blok o masie m, promieniu r i momencie bezwładności I przerzucono lekką nić obciążając

ją na końcach odpowiednio masami m_1 i m_2. Zakładając, że nić nie ślizga się po bloku i że tarcie można zaniedbać, obliczyć siły naprężające nici T_1 i T_2 oraz siłę F działającą w punkcie zawieszenia bloku.

Odp. $T_1 = \frac{2m_1 m_2 g + m_1 g I/r^2}{m_1 + m_2 + I/r^2}$, $T_2 = \frac{2m_1 m_2 g + m_2 g I/r^2}{m_1 + m_2 + I/r^2}$, $F = T_1 + T_2 + mg$.

18. W warunkach omówionych w treści zad. 17 obliczyć energie kinetyczne mas m_1 i m_2, bloku oraz całkowitą energię kinetyczną układu, gdy masa m_1 obniży się o h względem położenia początkowego.

Odp. $E_1 = \frac{m_1(m_1 - m_2)gh}{m_1 + m_2 + I/r^2}$, $E_2 = \frac{m_2(m_1 - m_2)gh}{m_1 + m_2 + I/r^2}$, $E_b = \frac{I(m_1 - m_2)g}{m_1 r^2 + m_2 r^2 + I}$, $E_c = (m_1 - m_2)gh$.

19. W warunkach omówionych w treści zad. 17 obliczyć przyspieszenie kątowe obrotu bloku.

Odp. $\frac{(m_1 - m_2)gr}{(m_1 + m_2)r^2 + I}$.

20. Koło zamachowe o masie $m = 30$ ton i ramieniu bezwładności $k = 5$ m robi 60 obrotów na minutę. Po zastosowaniu siły hamującej koło się zatrzymało w ciągu 6,3 minuty. Jak wielki był moment siły hamującej? Zakładamy stałość tego momentu.

Odp. $12{,}5 \cdot 10^3$ N · m.

21. Jednorodna kula o promieniu r stacza się pod wpływem własnego ciężaru po równi pochyłej nachylonej pod kątem φ względem poziomu. Jaka będzie prędkość środka masy tej kuli po przebyciu drogi s? Czy prędkość środka masy tej kuli byłaby mniejsza czy większa, gdyby kula zsuwała się po tej samej równi bez tarcia?

Odp. $\sqrt{\frac{gs\sin\varphi}{0{,}7}}$.

22. Łyżwiarz wykonujący piruet obraca się dokoła osi pionowej wykonując 2 obroty na sekundę. Przez rozsunięcie rąk zwiększył on swój moment bezwładności z 2 kg · m² na 2,1 kg · m². Jakiej zmianie uległa przy tym prędkość kątowa?

Odp. 0,6 rad/s.

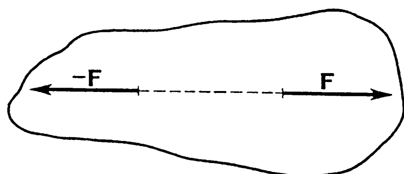
ROZDZIAŁ 6

Elementy statyki ciała sztywnego

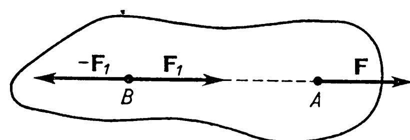
6.1. Postulat statyki

W rozdziale niniejszym będziemy się zajmować warunkami równoważenia się sił działających na ciało sztywne.

Przyjmujemy bez dowodu, jako postulat statyki, *że dwie siły równe, działające na ciało sztywne wzdłuż tej samej prostej, lecz mające zwroty przeciwne, równoważą się wzajemnie* (rys. 6.1).



Rys. 6.1



Rys. 6.2

Opierając się na tym postulacie można udowodnić możliwość przesuwania punktu przyłożenia siły działającej na ciało sztywne wzdłuż prostej jej działania w obrębie ciała sztywnego. Innymi słowy, można udowodnić, że *siła działająca na ciało sztywne jest wektorem przesuwalnym*. Załóżmy, że na punkt A ciała sztywnego (rys. 6.2) działa siła **F**. Na prostej działania tej siły obieramy punkt B i odkładamy dwie siły $\mathbf{F}_1$ i $-\mathbf{F}_1$ co do wartości równe F. Wolno tak postąpić, gdyż zgodnie z postulatem statyki takie postępowanie nie wywołuje żadnych zmian w stanie ruchu lub spoczynku ciała sztywnego. Z trzech sił działających obecnie na ciało sztywne bierzemy pod uwagę siłę **F** i $-\mathbf{F}_1$. Według postulatu te dwie siły równoważą się wzajemnie. Możemy więc ich nie uwzględniać. Pozostaje tylko siła $\mathbf{F}_1$ równa sile **F**, lecz przesunięta do innego punktu na prostej działania siły **F**.

Badanie zagadnienia równowagi ciała sztywnego pod działaniem większej liczby sił rozpoczniemy od tego, że będziemy szukali jednej siły zastępującej w działaniu wszystkie siły dane. Tę siłę będziemy nazywali *wypadkową sił danych*. Jeśli działanie

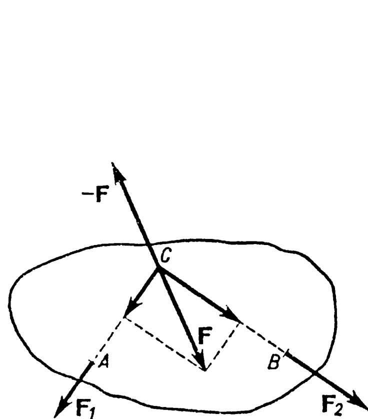
wielu sił da się sprowadzić do jednej siły wypadkowej, to zgodnie z postulatem statyki *siła równoważąca* wszystkie siły dane będzie równa ich wypadkowej, będzie skierowana wzdłuż tej samej prostej, lecz w stronę przeciwną.

Jeśli działanie kilku sił nie może być sprowadzone do działania jednej siły wypadkowej, to warunek równowagi staje się nieco bardziej skomplikowany. Omówimy go w § 3 tego rozdziału.

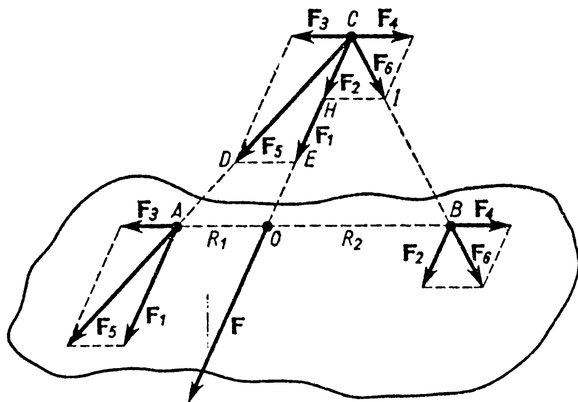
6.2. Przegląd różnych przypadków działania sił na ciało sztywne

Przegląd działania sił na ciało sztywne przeprowadzimy rozważając systematycznie następujące przypadki:

1. *Działanie sił leżących w jednej płaszczyźnie.*
   a) siły nierównoległe względem siebie,
   b) siły równoległe o zwrotach zgodnych,
   c) siły równoległe o zwrotach przeciwnych, nierówne co do wartości.
   d) siły równoległe o zwrotach przeciwnych, równe co do wartości.
2. *Działanie sił nie leżących w jednej płaszczyźnie.*



Rys. 6.3



Rys. 6.4

1a. Siły $\mathbf{F}_1$ i $\mathbf{F}_2$ działające na punkty A i B ciała sztywnego (rys. 6.3) przenosimy bez zmiany kierunku i wartości do punktu przecięcia prostych ich działania, a więc do punktu C. Budujemy na nich równoległobok. Przekątna $\mathbf{F}$ jest wypadkową obu sił. Do zrównoważenia sił $\mathbf{F}_1$ i $\mathbf{F}_2$ wystarcza siła $-\mathbf{F}$ równa wypadkowej, działająca wzdłuż tej samej prostej, lecz mająca przeciwny zwrot.

Gdy punkt C leży poza ciałem, przy wyznaczaniu wypadkowej postępujemy podobnie, a następnie przesuwamy ją tak, by punkt jej przyłożenia znalazł się w obrębie ciała.

1b. Siłami pierwotnymi działającymi na ciało sztywne są siły równoległe o zwrotach zgodnych $\mathbf{F}_1$ i $\mathbf{F}_2$ (rys. 6.4). Chcemy ten przypadek sprowadzić do omówionego wyżej działania sił przecinających się. W tym celu odkładamy w punktach A i B dwie do-

datkowe siły $\mathbf{F}_3$ i $\mathbf{F}_4$ równoważące się w działaniu, gdyż siła $\mathbf{F}_3 = -\mathbf{F}_4$. Znajdujemy na podstawie zasady równoległoboku wypadkowe sił $\mathbf{F}_1$ i $\mathbf{F}_3$ oraz $\mathbf{F}_2$ i $\mathbf{F}_4$. Są to siły $\mathbf{F}_5$ i $\mathbf{F}_6$. Przedłużamy ich kierunki aż do przecięcia w punkcie C i przenosimy do tego punktu siły $\mathbf{F}_5$ i $\mathbf{F}_6$. Rozkładamy je na składowe $\mathbf{F}_3$ i $\mathbf{F}_1$ oraz $\mathbf{F}_4$ i $\mathbf{F}_2$. Siłami $\mathbf{F}_3$ i $\mathbf{F}_4$ więcej się już nie zajmujemy, ponieważ równoważą się one wzajemnie. Okazuje się, że otrzymane w ten sposób siły składowe $\mathbf{F}_1$ i $\mathbf{F}_2$ działają wzdłuż tej samej prostej w tym samym kierunku. Ich wypadkowa równa się więc ich sumie arytmetycznej:

$$F = F_1 + F_2.$$

Siła wypadkowa jest równoległa do sił danych i skierowana z nimi zgodnie. Siłę wypadkową $\mathbf{F}$ można przyłożyć do ciała sztywnego w dowolnym punkcie prostej przechodzącej przez C i równoległej do kierunku sił pierwotnych $\mathbf{F}_1$ i $\mathbf{F}_2$.

Ustalmy warunek, jaki spełniać musi każdy z punktów tej prostej, np. punkt O leżący na odcinku AB łączącym punkty przyłożenia sił $\mathbf{F}_1$ i $\mathbf{F}_2$. Oznaczmy

$$AO = R_1, \quad OB = R_2.$$

Z podobieństwa trójkątów AOC i CDE oraz COB i CHI wynikają proporcje

$$\frac{R_1}{DE} = \frac{CO}{F_1} \quad \text{oraz} \quad \frac{R_2}{HI} = \frac{CO}{F_2}.$$

Uwzględniając równość DE i HI otrzymujemy

$$R_1 F_1 = R_2 F_2. \tag{6.1}$$

Jak łatwo sprawdzić, równanie to odpowiada innemu, a mianowicie

$$-\mathbf{R}_1 \times \mathbf{F}_1 = \mathbf{R}_2 \times \mathbf{F}_2, \tag{6.2}$$

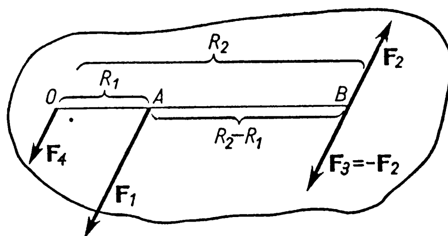
gdyż wobec równości sinusów kątów między $\mathbf{R}_1$ i $\mathbf{F}_1$ oraz $\mathbf{R}_2$ i $\mathbf{F}_2$ równanie (6.2) sprowadza się do równania (6.1). Wyrażenia po obu stronach równania (6.2) przedstawiają momenty odpowiednich sił względem punktu O. Zwroty obu momentów sił są przeciwne: uwzględniamy to wprowadzając przy jednym z iloczynów wektorowych znak minus. Zgodnie z ogólnie przyjętą umową momentom sił wywołującym obrót ciała w kierunku zgodnym z ruchem wskazówek zegara przypisuje się znak plus i odwrotnie. Łatwo sprawdzić, że w rozważanym przypadku moment siły $\mathbf{F}_1$ względem punktu O jest ujemny, zaś moment siły $\mathbf{F}_2$ — dodatni. Wobec równoczesnego działania obu sił i równych wartości liczbowych ich momentów względem punktu O suma ich momentów, czyli wypadkowy moment względem punktu O równa się zeru. Można udowodnić, że to samo obowiązuje względem dowolnego punktu leżącego na prostej działania wypadkowej.

Z równania (6.1) wynika proporcja

$$R_1 : R_2 = F_2 : F_1. \tag{6.3}$$

Innymi słowy, prosta działania siły wypadkowej dzieli odległość między punktami przyłożenia dwóch danych sił równoległych na odcinki odwrotnie proporcjonalne do wartości tych sił.

Doszliśmy do wniosku, że *siły równoległe zgodne można zastąpić jedną siłą wypadkową równoległą do sił danych, skierowaną zgodnie z nimi, równą ich sumie arytmetycznej i przyłożoną w takim punkcie, względem którego momenty obu sił danych są liczbowo równe.* Do zrównoważenia sił równoległych zgodnych wystarcza więc jedna siła równa wypadkowej, lecz przeciwnie skierowana.



Rys. 6.5

1c. Działanie sił nierównych równoległych, mających zwroty przeciwne. Niech na ciało sztywne działają siły $\mathbf{F}_1$ i $\mathbf{F}_2$ (rys. 6.5). Potraktujmy większą z nich, np. $\mathbf{F}_1$, jako wypadkową dwóch sił równoległych zgodnych. Jedna ze składowych $\mathbf{F}_3$ niech będzie równa $-\mathbf{F}_2$ i niech będzie przyłożona w tym samym punkcie B co siła $\mathbf{F}_2$. Wobec tego druga składowa $\mathbf{F}_4$ musi być równa liczbowo F_1-F_2. Punkt jej przyłożenia musi być tak dobrany, aby było spełnione równanie

$$OA \cdot (F_1-F_2) = AB \cdot F_2. \tag{6.4}$$

Zobaczmy, do czego sprowadza się działanie sił pierwotnych $\mathbf{F}_1$ i $\mathbf{F}_2$. Siłę $\mathbf{F}_1$ możemy zaniedbać, gdyż zastąpiliśmy ją siłami składowymi $\mathbf{F}_3$ i $\mathbf{F}_4$. Ale siła $\mathbf{F}_3$ zgodnie z postulatem statyki równoważy się z siłą $\mathbf{F}_2$. Możemy więc i te dwie siły zaniedbać. Zostaje tylko siła $\mathbf{F}_4$. Wyciągamy wniosek, że *dwie siły równoległe nierówne, przeciwnie skierowane, można zastąpić jedną siłą wypadkową do nich równoległą, równą ich różnicy arytmetycznej, mającą zwrot zgodny z siłą większą.*

Aby określić punkt przyłożenia tej wypadkowej, wróćmy do równania (6.4). Oznaczmy $OA = R_1$, $OB = R_2$. W takim razie $AB = R_2-R_1$ i równanie (6.4) przyjmuje postać

$$R_1(F_1-F_2) = (R_2-R_1)F_2,$$

skąd

$$\frac{F_1-F_2}{F_2} = \frac{R_2-R_1}{R_1},$$

czyli

$$\frac{F_1}{F_2} = \frac{R_2}{R_1} \quad \text{lub} \quad F_1R_1 = F_2R_2.$$

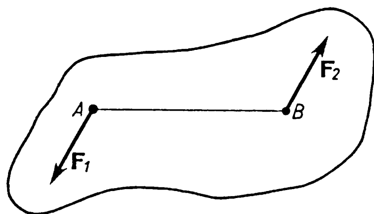
Punkt O przyłożenia siły wypadkowej leżący na przedłużeniu odcinka AB łączącego punkty przyłożenia obu sił danych $\mathbf{F}_1$ i $\mathbf{F}_2$, ma tę własność, że jego odległości od punktów A i B są odwrotnie proporcjonalne do wartości obu sił. Leży on zatem po stronie siły większej.

Inaczej ujmując ten sam warunek powiemy, że momenty obu sił danych względem punktu O są liczbowo równe, a przeciwnie skierowane:

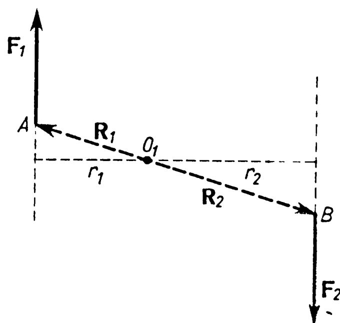
$$\mathbf{R}_1 \times \mathbf{F}_1 = -\mathbf{R}_2 \times \mathbf{F}_2 .$$

Ta równość momentów sił obowiązuje względem każdego punktu leżącego na kierunku wypadkowej. Moment wypadkowy obu sił względem punktu O równa się zeru.

Podobnie jak w poprzednich przypadkach, *siły równoległe nierówne, przeciwnie skierowane, można zrównoważyć jedną siłą, równą ich wypadkowej, działającą wzdłuż tej samej prostej co wypadkowa, lecz w przeciwnym kierunku.* Zwrot siły równoważącej jest więc zgodny z mniejszą z danych sił.



Rys. 6.6



Rys. 6.7

1d. Na ciało sztywne działają dwie siły równe, równoległe, przeciwnie skierowane $\mathbf{F}_1$ i $\mathbf{F}_2$ (rys. 6.6). Różnica wartości tych sił równa się zeru, a więc siły *te mają wypadkową równą zeru. Nie można ich ani zastąpić, ani zrównoważyć pojedynczą siłą.* W tym przypadku mówimy o działaniu *pary sił.*

Suma momentów obu sił pary względem dowolnego punktu nie równa się zeru. Sprawdzimy to obierając kolejno punkt odniesienia między siłami pary i poza nimi. W przypadku pierwszym (rys. 6.7) momenty sił $\mathbf{F}_1$ i $\mathbf{F}_2$ mają te same znaki, oba są dodatnie; obie siły wywołują obrót dokoła O_1 w kierunku zgodnym z ruchem wskazówek zegara. Zgodnie z oznaczeniami na rysunku suma momentów wynosi

$$r_1 F_1 + r_2 F_2 = F(r_1 + r_2),$$

gdyż

$$F_1 = F_2 = F.$$

Odległość prostych dzialania obu sił pary nazywamy ramieniem pary i oznaczamy literą d:

$$d = r_1 + r_2 .$$

Para sił ma zatem moment wypadkowy względem punktu O_1 równy Fd, czyli równy iloczynowi jednej z sił pary przez ramię pary. Ten moment wypadkowy nazywamy *momentem pary.*

Obierzmy punkt O_2 poza obszarem ograniczonym przez kierunki sił pary (6.8). Wtedy oba momenty sił względem punktu O_2 mają znaki przeciwne. Moment siły $\mathbf{F}_1$

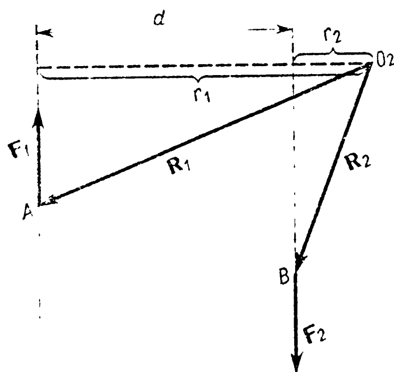
jest dodatni i równy r_1F_1. Moment siły $\mathbf{F}_2$ jest ujemny i równy $-r_2F_2$. Suma obu momentów względem punktu O_2 wynosi

$$r_1F_1+(-r_2F_2) = (d+r_2)F-r_2F = Fd.$$

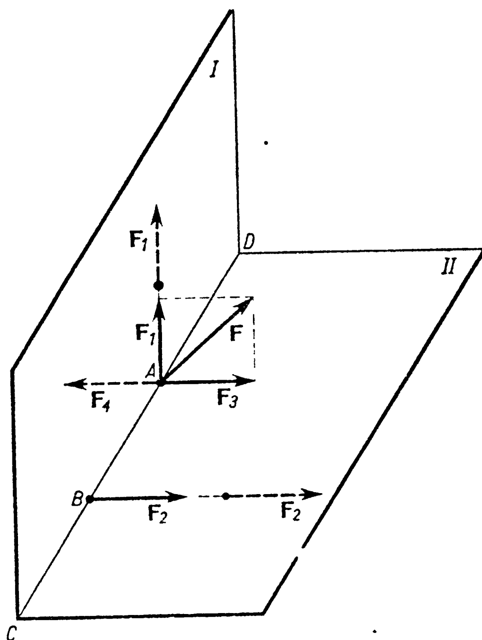
Jak widać, *moment wypadkowy pary* nie zależy od wyboru punktu, względem którego moment obliczamy. *Równa się on zawsze iloczynowi wartości jednej z sił pary i ramienia pary sił*:

$$M_{\text{pary sił}} = Fd. \tag{6,5}$$

Para sił zawsze usiłuje wywołać obrót ciała. Para taka jak na rys. 6.8 wywołuje obrót w kierunku zgodnym z ruchem wskazówek zegara. Pod działaniem pojedynczej pary sił ciało sztywne nie pozostaje w spoczynku, mimo że wypadkowa obu sił pary równa się zeru. Do zrównoważenia pary sił potrzebna jest druga para o momencie takim samym co do wartości, lecz przeciwnym co do znaku.



Rys. 6.8



Rys. 6.9

2. Przechodzimy obecnie do przypadku, gdy siły $\mathbf{F}_1$ i $\mathbf{F}_2$ leżą w różnych (dowolnych) płaszczyznach przecinających się wzdłuż prostej CD (rys. 6.9). Ponieważ siły w ciele sztywnym wolno przenosić wzdłuż prostych ich działania, więc przenosimy je tak, by ich punkty przyłożenia leżały na prostej CD. Taki stan rzeczy przedstawia rys. 6.9. W punkcie A przyłożenia siły $\mathbf{F}_1$ odłóżmy dodatkowo dwie siły równe i przeciwnie skierowane $\mathbf{F}_3$ i $\mathbf{F}_4$, równe co do wartości sile $\mathbf{F}_2$. Obie te siły leżą w płaszczyźnie II. Siły $\mathbf{F}_1$ i $\mathbf{F}_3$ działające na ten sam punkt mają wypadkową $\mathbf{F}$. Siły $\mathbf{F}_2$ i $\mathbf{F}_4$ leżą obie w płaszczyźnie II, są równe, równoległe, mają zwroty przeciwne, a więc stanowią parę.

Innymi słowy, działania dwóch sił nie leżących w jednej płaszczyźnie nie można sprowadzić do działania jednej siły wypadkowej. Nie można więc też zrównoważyć ich jedną siłą. *Działanie dwóch sił nie leżących w jednej płaszczyźnie sprowadza się do jednej siły wypadkowej i do pary sił.* Podobnie działanie większej liczby dowolnie zorientowanych sił sprowadza się, ogólnie biorąc, do działania jednej siły wypadkowej i do jednej pary sił.

6.3. Warunek równowagi statycznej ciała sztywnego

Przez *warunek równowagi statycznej ciała sztywnego* rozumiemy warunek, który musi być spełniony, aby ciało pierwotnie spoczywające pozostawało nadal w spoczynku. Będziemy go dalej nazywali krótko *warunkiem równowagi ciała sztywnego.*

Po rozważeniu punktów 1a, 1b, 1c bylibyśmy skłonni powiedzieć, że warunkiem równowagi ciała jest to, by wypadkowa wszystkich sił działających równała się zeru. Doszliśmy jednak do punktu 1d i stwierdziliśmy, że pod działaniem pary sił nie mamy równowagi, mimo że wypadkowa siła równa się zeru. Stąd wniosek, *że warunek, aby wypadkowa sił działających na ciało była równa zeru, jest warunkiem koniecznym, ale niewystarczającym do utrzymania równowagi ciała.* Równocześnie musi być spełniony drugi warunek, a mianowicie: *wypadkowy moment wszystkich sił działających względem dowolnego punktu w przestrzeni musi być równy zeru.*

Nasuwa się pytanie, czy zastępując siły przecinające się lub siły równoległe, nie tworzące pary, jedną siłą wypadkową nie zaniedbaliśmy przy tym istnienia momentu tych sił. Może siły pierwotne mają inny moment względem dowolnego punktu przestrzeni niż zastępująca je w działaniu ich wypadkowa. Zamiast przeprowadzać dość skomplikowany ogólny dowód matematyczny udowodnimy, że suma momentów sił składowych równa się momentowi wypadkowej dla dowolnego punktu wybranego na prostej działania wypadkowej. Z określenia momentu siły bezpośrednio wynika, że moment wypadkowej względem dowolnego punktu leżącego na prostej jej działania równa się zeru. Wystarczy zatem udowodnić, że suma momentów sił składowych względem dowolnego punktu wypadkowej równa się zeru.

Rysunek 6.10 odnosi się do działania sił nierównoległych. Moment siły $\mathbf{F}_1$ względem punktu O liczbowo równa się r_1F_1 i jest dodatni. Moment siły $\mathbf{F}_2$ liczbowo równa się r_2F_2 i jest ujemny. Chcemy udowodnić, że wartości liczbowe obu momentów są jednakowe. Opieramy się na równości pól trójkątów ABC i ABD:

$$\frac{h_1F_1}{2} = \frac{h_2F_2}{2},$$

skąd

$$\frac{F_1}{F_2} = \frac{h_2}{h_1}. \tag{6.6}$$

Równocześnie z podobieństwa dwóch par trójkątów, a mianowicie AKO i AGB oraz ALO i ABE wynikają następujące proporcje:

$$\frac{r_1}{h_1} = \frac{AO}{AB}, \quad \frac{r_2}{h_2} = \frac{AO}{AB}.$$

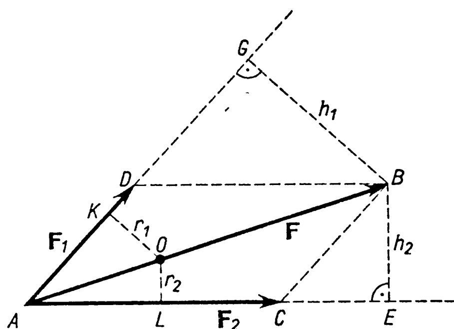
Stąd otrzymujemy

$$\frac{r_1}{h_1} = \frac{r_2}{h_2}, \quad \text{czyli} \quad \frac{h_2}{h_1} = \frac{r_2}{r_1}.$$

Podstawiając do wzoru (6.6) mamy

$$\frac{F_1}{F_2} = \frac{r_2}{r_1}, \quad \text{czyli} \quad r_1 F_1 = r_2 F_2.$$

Udowodniliśmy, że liczbowe wartości obu momentów sił względem punktu O są jednakowe. Wobec ich przeciwnych znaków suma ich równa się zeru. To samo stwierdziliśmy poprzednio dla sił równoległych zgodnych i przeciwnych, nie tworzących pary.



Rys. 6.10

Uogólniając ten wynik powiemy, że *suma momentów sił składowych równa się momentowi wypadkowej względem każdego, dowolnego punktu przestrzeni*. Możemy zatem rzeczywiście zastępować siły przecinające się lub równoległe nie tworzące pary jedną siłą, ich wypadkową. Wprowadzając dodatkowo siłę równoważącą sprowadzimy do zera zarówno ostateczną wypadkową, jak i wypadkowy moment wszystkich sił, gdyż jak pamiętamy, z rozumowania opartego na analizie rys. 5.4 wynika, że w odniesieniu do dowolnego punktu przestrzeni wypadkowy moment dwóch sił równych sobie, działających wzdłuż tej samej prostej i mających zwroty przeciwne, równa się zeru.

Pojedyncza para sił ma wypadkową równą zeru, lecz jej moment nie równa się zeru. *Dla zrównoważenia momentu pary konieczne jest użycie drugiej pary sił o momencie tym samym co do wartości, lecz przeciwnie skierowanym*. Wtedy moment wypadkowy obu par równa się zeru. Para sił **Q** i —**Q** (rys. 6.11) użyta do zrównoważenia pary danej **F** i —**F** może się składać z sił o innej wielkości, muszą one jednak być tak rozmieszczone, aby

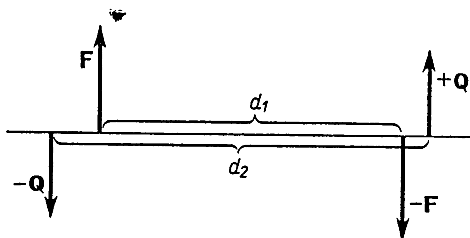
$$Fd_1 = Qd_2,$$

gdzie d_1 i d_2 stanowią ramiona odpowiednio pary **F** i —**F** oraz **Q** i —**Q**. Para sił **Q** i —**Q** wywołuje obrót w kierunku przeciwnym do kierunku ruchu wskazówek zegara (rys. 6.11), a zatem ma moment ujemny.

Z powyższych rozważań wynika, że ogólny warunek równowagi ciała sztywnego poddanego działaniu dowolnej liczby n sił dowolnie rozmieszczonych można matematycznie ująć następująco:

$$\sum_{i=1}^{n} \mathbf{F}_i = 0 \text{ — } \textit{wypadkowa wszystkich sił równa się zeru,} \tag{6.7}$$

$$\sum_{i=1}^{n} \mathbf{M}_i = 0 \text{ — } \textit{wypadkowy moment wszystkich sił względem dowolnego punktu w przestrzeni równa się zeru.} \tag{6.8}$$



Rys. 6.11

Równanie wektorowe (6.7) jest równoważne następującym trzem równaniom skalarnym:

$$\begin{aligned} F_x &= F_{1x}+F_{2x}+F_{3x}+ \ldots +F_{nx} = 0,\\ F_y &= F_{1y}+F_{2y}+F_{3y}+ \ldots +F_{ny} = 0,\\ F_z &= F_{1z}+F_{2z}+F_{3z}+ \ldots +F_{nz} = 0. \end{aligned} \tag{6.9a}$$

Z równań tych wynika, że *w przypadku równowagi statycznej suma rzutów wszystkich sił działających na ciało sztywne na każdą z trzech dowolnie wybranych wzajemnie prostopadłych osi współrzędnych xyz równa się zeru.*

Podobnie równanie wektorowe (6.8) jest równoważne układowi równań skalarnych:

$$\begin{aligned} M_x &= M_{1x}+M_{2x}+M_{3x}+ \ldots +M_{nx} = 0,\\ M_y &= M_{1y}+M_{2y}+M_{3y}+ \ldots +M_{ny} = 0,\\ M_z &= M_{1z}+M_{2z}+M_{3z}+ \ldots +M_{nz} = 0. \end{aligned} \tag{6.9b}$$

Z równań tych wynika, że w *przypadku równowagi statycznej suma rzutów wszystkich momentów sił działających na ciało sztywne na każdą z trzech dowolnie wybranych wzajemnie prostopadłych osi jest równa zeru.*

Ostatecznie zatem ogólny warunek równowagi ciała sztywnego sprowadza się do sześciu równań skalarnych.

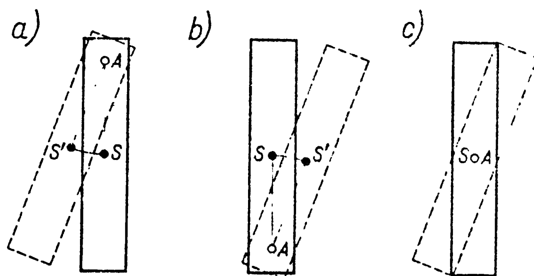
Gdy ciało sztywne ma ograniczoną swobodę poruszania się w przestrzeni, np. może się poruszać tylko po jednej płaszczyźnie lub po określonej krzywej, wtedy liczba równań określających warunek równowagi statycznej znacznie się redukuje. W przypadku

ruchu ograniczonego do jednej płaszczyzny suma rzutów wszystkich sił działających na każdą z dwóch dowolnie wybranych prostopadłych do siebie osi leżących w tej płaszczyźnie, jak również suma momentów sił względem dowolnie wybranej osi prostopadłej do tej płaszczyzny, musi być równa zeru. Zwracamy uwagę na ten szczególny przypadek, gdyż jest on często spotykany w praktyce. Rozwiązując np. zadania, których treść wymaga zastosowania warunków równowagi w odniesieniu do sił działających w jednej płaszczyźnie należy 1) tak wybrać dwie prostopadłe osie leżące w tej płaszczyźnie, by obliczanie rzutów działających sił na wybrane kierunki było możliwie najprostsze, 2) tak wybrać oś prostopadłą do danej płaszczyzny, by możliwie maksymalna liczba momentów sił względem tej osi była równa zeru. Wygodny jest wybór takiej osi, by możliwie jak najwięcej kierunków działania sił nieznanych tę oś przecinało, bowiem ich momenty względem takiej osi są równe zeru.

6.4. Równowaga ciał w polu grawitacyjnym

Przeanalizujemy zagadnienie równowagi ciał w polu grawitacyjnym, biorąc pod uwagę ciała zawieszone.

Opierając się na ogólnym warunku równowagi brył łatwo stwierdzimy, że ciało zawieszone może być w równowadze jedynie wtedy, gdy środek ciężkości S ciała leży na linii pionowej przechodzącej przez punkt zawieszenia. Tylko bowiem w takim przypadku wypadkowy moment siły ciężkości i reakcji zawieszenia będzie równy zeru. Łatwo jednak sprawdzić, że umieszczenie punktu zawieszenia A powyżej, poniżej lub w samym środku ciężkości (rys. 6.12a, b, c) ma zasadniczy wpływ na *rodzaj równowagi.* W przypadku pierwszym (rys. 6.12a) mówimy o *równowadze trwałej,* w drugim (rys. 6.12b) — o *równowadze chwiejnej,* w trzecim (rys. 6.12c) — o *równowadze obojętnej.*



Rys. 6.12

Pewne charakterystyczne cechy wspomnianych trzech rodzajów równowagi można zaobserwować wychylając lekko ciało od rozważanego położenia i pozostawiając je następnie samemu sobie. Okazuje się, że przy takim odchyleniu ciała od położenia równowagi trwałej do położenia sąsiedniego występuje przesunięcie się środka ciężkości do położenia S', wyższego niż S (rys. 6.12a). Pojawia się przy tym moment obracający ciało w stronę pierwotnego położenia: ciało nie podtrzymywane wraca do tego położenia.

Odchylenie ciała od położenia równowagi chwiejnej do położenia sąsiedniego powoduje przesunięcie środka ciężkości do punktu S', leżącego poniżej punktu S (rys. 6.12b). Znów pojawia się moment obracający, tym razem jednak nie w stronę pierwotnego położenia, lecz do jakiegoś nowego położenia równowagi.

Odchylenie ciała od położenia równowagi obojętnej do położenia sąsiedniego nie ma wpływu na wysokość środka ciężkości. Nowe położenie nie różni się — z punktu widzenia warunków równowagi — od położenia pierwotnego. Ciało zachowuje je jako nowe położenie równowagi obojętnej.

Tak więc w polu grawitacyjnym w położeniu równowagi trwałej wysokość środka ciężkości jest mniejsza niż w położeniach sąsiednich ciała, w położeniu równowagi chwiejnej — większa niż w położeniach sąsiednich, a w położeniu równowagi obojętnej — taka sama jak w położeniach sąsiednich. Zmiany wysokości środka ciężkości przy małym odchyleniu od badanego położenia są jednym z kryteriów, według których można określać rodzaj równowagi w polu grawitacyjnym. Odsyłamy Czytelnika do pytania 9, nawiązującego do przykładu zastosowania tego kryterium do ciał spoczywających na podstawie poziomej.

6.5. Związek rodzajów równowagi z energią potencjalną

Drugie kryterium, na podstawie którego można rozróżniać rodzaje równowagi, wiąże się z energią potencjalną. Łatwo stwierdzić, że skoro *równowagę trwałą* w polu grawitacyjnym charakteryzuje najniższe położenie środka ciężkości ciała, to wartość energii potencjalnej grawitacyjnej, związanej z tym rodzajem równowagi, jest najmniejsza (*minimum energii potencjalnej*).

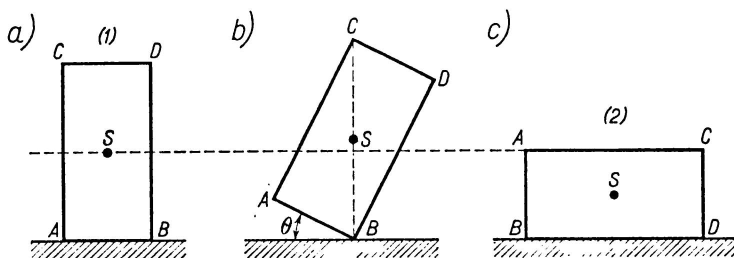
Odwrotna sytuacja energetyczna charakteryzuje *równowagę chwiejną*: z najwyższym położeniem środka ciężkości wiąże się największa wartość energii potencjalnej grawitacyjnej (*maksimum energii potencjalnej*).

Równowagę obojętną charakteryzuje nie zmieniająca się wysokość środka ciężkości i *stała wartość energii* potencjalnej grawitacyjnej.

Jak wynika z poprzednio rozważanego przykładu, ciało wytrącone z położenia równowagi trwałej i pozostawione samemu sobie wraca do stanu początkowego, a więc do stanu o minimalnej energii potencjalnej. Jest to normalny przebieg procesów samorzutnych w przyrodzie. Zilustrować go można przykładem naciągniętej lub ściśniętej sprężyny, która po zwolnieniu z „więzów" samorzutnie przechodzi do stanu odpowiadającego minimum energii potencjalnej sprężystej.

Na zakończenie zwrócimy jeszcze uwagę na zmiany E_p, związane z przejściem badanego układu od jednego stanu równowagi trwałej do innego. Weźmy np. pod uwagę prostopadłościan, który przez obrót dookoła jednej z krawędzi przeprowadzamy z położenia przedstawionego na rys. 6.13a do położenia przedstawionego na rys. 6.13c, stopniowo zwiększając kąt odchylenia θ. Rys. 6.13b przedstawia jedno z położeń pośrednich, a mianowicie takie, przy którym środek ciężkości S ma największą wysokość. Jest to więc położenie równowagi chwiejnej. Łatwo można stwierdzić, że aczkolwiek w stanie końcowym (2) energia potencjalna jest mniejsza niż w stanie początkowym (1),

to jednak przejście od stanu (1) do stanu (2) nie mogło się odbyć samorzutnie, gdyż oba te stany są przedzielone stanem równowagi chwiejnej o energii potencjalnej $E_p^{(ch)}$, większej zarówno od $E_p^{(1)}$, jak i od $E_p^{(2)}$. Między tymi stanami istnieje zatem zapora energetyczna — *bariera energii potencjalnej*. Chcąc wywołać przejście od stanu (1) do (2), trzeba układowi dostarczyć energię (np. w postaci pracy) w ilości równej różnicy energii potencjalnej układu w stanie równowagi chwiejnej i w stanie (1).



Rys. 6.13

Pytania. Czy bariera energii potencjalnej przy przejściu od stanu (1) do stanu (2) jest taka sama jak przy przejściu od stanu (2) do stanu (1)?

Czy słuszne byłoby nazywanie bariery energii potencjalnej $E_p^{(ch)}-E_p^{(1)}$ energią aktywacji procesu przejścia od stanu (1) do stanu (2)?

6.6. Wagi laboratoryjne

Przypomnimy kilka zasadniczych wiadomości o dźwigniach, aby łatwiej zrozumieć działanie wagi, jednego z podstawowych przyrządów pomiarowych.

Dźwignia jest to ciało sztywne podparte w ten sposób, że może się obracać dokoła osi podparcia. Jeżeli siły działające na dźwignię są rozmieszczone po obu stronach osi podparcia, mówimy o *dźwigni dwustronnej*. Jeśli ramiona sił działających na dźwignię są jednakowe, to *dźwignia* jest *równoramienna*.

Warunek równowagi dźwigni jest taki sam, jak każdej bryły sztywnej: wypadkowa siła i wypadkowy moment siły względem dowolnego punktu (a więc i względem punktu podparcia) muszą być równe zeru.

Rysunek 6.14 przedstawia równowagę najprostszej dwustronnej równoramiennej dźwigni. Równe siły $\mathbf{F}_1$ i $\mathbf{F}_2$ mają wypadkową $\mathbf{F}$, przyczepioną w środku między nimi w punkcie O. Siła wypadkowa jest zrównoważona przez siłę reakcji podparcia $\mathbf{R}$, a więc

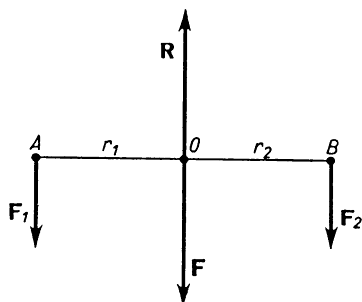
$$F_1+F_2+(-R)=0.$$

Momenty sił $\mathbf{F}_1$ i $\mathbf{F}_2$ względem punktu O są liczbowo równe sobie, $r_1F_1=r_2F_2$, lecz mają znaki przeciwne, gdyż obie siły usiłują wywołać obrót dokoła punktu O w przeciwnych kierunkach, a zatem

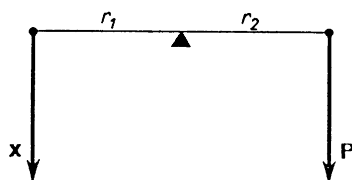
$$-r_1F_1+r_2F_2=0.$$

Waga laboratoryjna techniczna czy analityczna przedstawia w zasadzie dźwignię dwu-

stronną równoramienną. Podstawową część wagi laboratoryjnej stanowi belka podparta w środku, o symetrycznej budowie w stosunku do osi podparcia. Na końcach belki zawieszone są szalki o jednakowym ciężarze. Punkty zawieszenia szalek, a równocześnie punkty działania sił na belkę są równoodległe od osi podparcia. Środek ciężkości belki i szalek przypada poniżej osi podparcia. Belka może wykonywać wahania dokoła osi podparcia. Położenie poziome belki przy szalkach nieobciążonych i przy szalkach obciążonych jednakowymi ciężarami jest położeniem równowagi trwałej. Środek ciężkości belki leży wtedy w płaszczyźnie pionowej, przechodzącej przez oś podparcia. Wypadkowa wszystkich sił łącznie z siłą reakcji podparcia i wypadkowy moment względem punktu podparcia równają się zeru.



Rys. 6.14



Rys. 6.15

Przy budowie wagi laboratoryjnej zwraca się bardzo wielką uwagę na uzyskanie symetrii budowy względem punktu podparcia. Oba ramiona powinny być jednakowe co do długości i ciężaru oraz obie szalki powinny być równie ciężkie. Jeśli ten warunek jest spełniony, to ten sam odważnik położony na jednej, potem na drugiej szalce wywoła to samo wychylenie belki. Wtedy mówimy, że *waga* jest *rzetelna*. Rzetelność wagi sprawdzamy w ten sposób, że przy ważeniu jakiegoś ciała doprowadzamy wagę do równowagi, następnie przenosimy ciało i odważniki z jednej szalki na drugą. Jeśli w dalszym ciągu istnieje równowaga, to waga jest rzetelna.

Na wadze nierzetelnej o nierównych ramionach można też dokładnie ważyć stosując metodę podwójnego ważenia albo metodę tarowania. W *metodzie podwójnego ważenia* ciało o szukanym ciężarze x ważymy dwukrotnie. Początkowo umieszczamy je np. na szalce lewej i równoważymy odważnikiem P. Wtedy (rys. 6.15)

$$xr_1 = Pr_2;$$

następnie przenosimy ciało na szalkę prawą i równoważymy odważnikiem Q:

$$xr_2 = Qr_1.$$

Mnożąc stronami oba równania i upraszczając przez $r_1 r_2$ otrzymujemy

$$x^2 = PQ,$$

skąd

$$x = \sqrt{PQ}.$$

Ciężar ciała ważonego na wadze nierzetelnej znajdujemy jako średnią geometryczną wartości odważników równoważących, uzyskanych przy podwójnym ważeniu.

W przypadku niewielkiej różnicy między P i Q można stosować wzór przybliżony:

$$x = \frac{P+Q}{2},$$

tzn. szukany ciężar ciała można uważać za średnią arytmetyczną ciężarów P i Q. Jak łatwo sprawdzić,

$$\sqrt{PQ} = \sqrt{\tfrac{1}{4}(P+Q)^2 - \tfrac{1}{4}(P-Q)^2}.$$

Ponieważ $P-Q$ jest wielkością małą, przeto $\frac{1}{4}(P-Q)^2$ możemy zaniedbać nie popełniając większego błędu. Stąd otrzymujemy wzór przybliżony:

$$\sqrt{PQ} \cong \tfrac{1}{2}(P+Q).$$

W metodzie tarowania ciało badane umieszczamy na jednej z szalek i równoważymy dowolnym ciałem sypkim, np. piaskiem. Następnie zdejmujemy ciało ważone i na jego miejsce umieszczamy taką liczbę odważników, aby zrównoważyły tarę, tzn. piasek. Ciężar odważników równa się wtedy ciężarowi ciała ważonego.

Zasadnicze równanie równowagi wagi rzetelnej

$$F_1 r_1 = F_2 r_2$$

sprowadza się do wzoru

$$m_1 g r_1 = m_2 g r_2,$$

skąd wobec równości r_1 i r_2 otrzymujemy

$$m_1 = m_2.$$

Z równowagi wagi równoramiennej można sądzić nie tylko o równości ciężarów, ale i o równości mas ciał umieszczonych na obu szalkach, gdyż wobec stosunkowo niewielkich rozmiarów geometrycznych wagi uzasadnione jest założenie jednakowej wartości g na obu szalkach*.

6.7. Gęstość a ciężar właściwy

Jak już podkreślaliśmy, za pomocą wagi można wyznaczać zarówno masy ciał, jak i ich ciężary. W układzie SI masy powinny być wyrażane w kilogramach (kg), natomiast ciężary — jako siły — w niutonach. Przypuszczalnie jednak jeszcze przez długi czas będą w użyciu jednostki ciężarowe siły.

Przez podzielenie całkowitej masy ciała jednorodnego lub całkowitego jego ciężaru przez objętość otrzymujemy dwie nowe wielkości mechaniczne, a mianowicie *gęstość* ϱ (zwaną czasem gęstością bezwzględną) i *ciężar właściwy* γ. Ich definicje odpowiadają równaniom

$$\varrho = m/V, \qquad \gamma = P/V; \tag{6.10}$$

---

* Pomijaliśmy w tych rozważaniach wpływ sił wyporu, działających na ciało ważone i odważniki.

gdzie m oznacza masę ciała, P — ciężar ciała, V — objętość. W odniesieniu do ciał niejednorodnych wzory (6.10) przedstawiają średnią gęstość i średni ciężar właściwy.

Jednostką gęstości w układzie SI jest kg/m^3 (por. tab. 6.1). *Jednostką ciężaru właściwego w układzie* SI jest N/m^3. Przypuszczalnie jednak jeszcze przez długi czas będą w użyciu jednostki G/cm^3 i kG/m^3 jako jednostki ciężaru właściwego. Należy jednak pamiętać, że nie są to jednostki spójne z jednostkami układu SI.

Tabela 6.1

Gęstości bezwzględne w kg/m^3 w temperaturze 18°C

Ciała stałe		Ciała ciekłe	
korek	~ $0{,}24 \cdot 10^3$	benzyna	~ $0{,}7 \cdot 10^3$
suche drewno sosnowe	~ $0{,}5 \cdot 10^3$	nafta	~ $0{,}75 \cdot 10^3$
suche drewno dębowe	~ $0{,}8 \cdot 10^3$	alkohol etylowy	~ $0{,}79 \cdot 10^3$
cegła	~ $1{,}5 \cdot 10^3$	woda (temp. 4°C)	~ $1{,}00 \cdot 10^3$
szkło	~ $2{,}5 \cdot 10^3$	rtęć	~ $13{,}546 \cdot 10^3$
glin	~ $2{,}67 \cdot 10^3$		
żelazo	~ $7{,}86 \cdot 10^3$	**Ciała gazowe w warunkach normalnych**	
stal	~ $7{,}8 \cdot 10^3$		
mosiądz	~ $8{,}4 \cdot 10^3$	wodór	~ 0,0899
nikiel	~ $8{,}8 \cdot 10^3$	azot	~ 1,250
miedź	~ $8{,}9 \cdot 10^3$	powietrze	~ 1,293
srebro	$10{,}5 \cdot 10^3$	tlen	~ 1,429
ołów	$11{,}3 \cdot 10^3$	dwutlenek węgla	~ 1,977
złoto	$19{,}3 \cdot 10^3$		
osm	$22{,}48 \cdot 10^3$		

6.8. Tarcie

Nawet w bardzo ogólnym ujęciu zagadnień z dziedziny kinematyki, dynamiki i statyki ciała sztywnego nie może zabraknąć podstawowych wiadomości o siłach, zwanych *siłami tarcia*, przeciwstawiających się ruchowi ciała stałego po powierzchni innego ciała stałego.

Od razu na wstępie musimy stwierdzić, że zjawisko występowania sił tarcia, zwane krótko tarciem, aczkolwiek bardzo pospolite, jest zjawiskiem dotychczas nie w pełni zbadanym. Podamy niżej kilka stwierdzonych doświadczalnie prawidłowości dotyczących tarcia przy przesuwaniu ciał, zwanego poślizgowym, i tarcia przy toczeniu, zwanego tocznym.

Doświadczenia wykazują, że *siła tarcia T występująca przy poślizgu ciała stałego po ciele stałym jest z nielicznymi wyjątkami proporcjonalna do siły F_n, przyciskającej ciało do podłoża.* Siłę F_n nazywamy czasem naciskiem:

$$T = \mu F_n, \tag{6.11}$$

gdzie μ oznacza *współczynnik tarcia poślizgowego*, zwanego też krótko *współczynnikiem tarcia*. Jest to liczba bezwymiarowa, wskazująca jaką część siły nacisku stanowi siła tarcia.

Drugą prawidłowością stwierdzoną doświadczalnie w odniesieniu do siły tarcia poślizgowego jest jej *niezależność od pola powierzchni zetknięcia poruszającego się ciała i podłoża*. Wartość współczynnika tarcia zależy w sposób zasadniczy od rodzaju stykających się powierzchni. Tak np. w przypadku powierzchni metalicznych istotną rolę odgrywa sposób ich obróbki, stopień zanieczyszczenia powierzchni i charakter zanieczyszczeń. Jako pewną ciekawostkę podamy, że współczynnik tarcia charakteryzujący dwie metaliczne płytki, stosunkowo duży przy znacznej „chropowatości" powierzchni, maleje przy ich wygładzaniu, ale tylko do pewnej granicy. Po bardzo dokładnym wygładzeniu powierzchni siła tarcia staje się bardzo duża: efekt jest taki, jak gdyby płytki zlepiały się ze sobą.

Efekt ostatnio omówiony, jak również wspomniane poprzednio prawidłowości charakteryzujące tarcie, spróbujemy wyjaśnić analizując nieco dokładniej mechanizm tarcia na przykładzie dwóch płytek metalicznych o polu powierzchni S. Na powierzchni każdej z płytek, nawet pozornie idealnie gładkiej, istnieją nierówności: wgłębienia i wypukłości. Po umieszczeniu jednej płytki na drugiej stykają się ze sobą tylko wierzchołki wypukłości. Pole S_0 powierzchni rzeczywistego styku jest znacznie (przeciętnie 10^4–10^5 razy) mniejsze od pola S powierzchni płytek. Stykające się ze sobą wypukłości obu płytek rozpłaszczają się tym bardziej, im większa jest siła nacisku F_n, niezależnie jednak od S.

Pole S_0, wynikające z rozpłaszczenia się wypukłości, można traktować jako sumę pól poszczególnych styków. Na każdym z nich, wobec małej odległości między cząsteczkami, działają duże siły przyciągania międzycząsteczkowego. Powstają w ten sposób więzy łączące obie płytki. Ruch względny płytek wymaga zrywania tych więzów, czyli stosowania zewnętrznej siły. Jeśli ruch jest jednostajny, to siła zewnętrzna równoważy opór przeciwstawiający się ruchowi, czyli właśnie siłę tarcia. Ale ta siła zewnętrzna (a zatem i siła tarcia) jest proporcjonalna do pola S_0, czyli proporcjonalna do nacisku F_n, a niezależna od S. Gwałtowny wzrost siły tarcia przy daleko posuniętym wygładzeniu powierzchni styku jest już zrozumiały bez dalszych wyjaśnień.

Jak już wspominaliśmy, istnieją odstępstwa od omówionych wyżej praw tarcia. Tak np. przy ruchu gumy po gumie siła tarcia maleje ze wzrostem siły F_n przyciskającej do podłoża.

Wśród praw doświadczalnych charakteryzujących tarcie czasem się wymienia prawo niezależności siły tarcia od prędkości ruchu względnego przesuwanych ciał. Trzeba jednak pamiętać, że to prawo jest spełnione tylko w ograniczonych zakresach zmian prędkości. Prędkościom znacznie się różniącym odpowiadają różne wartości współczynników tarcia (przy pozostałych warunkach nie zmienionych). Różnice wiążą się ze zmianą mechanizmu tarcia. Tak np. przy dużych prędkościach wchodzą w grę silne efekty temperaturowe, mające wpływ na właściwości powierzchni ciał, powodujące nawet czasem topienie na powierzchni i zmniejszające wielokrotnie współczynnik tarcia.

Mówiąc o tarciu należy rozróżniać *siłę tarcia kinetycznego*, występującego podczas ruchu ciała, i *siłę tarcia statycznego*, występującego na początku ruchu, tzn. przy przejściu ze stanu spoczynku do stanu ruchu. Siła tarcia statycznego jest również

proporcjonalna do siły nacisku F_n, lecz współczynnik proporcjonalności μ_0, zwany *współczynnikiem tarcia statycznego*, jest zwykle o 10–20% większy od współczynnika tarcia kinetycznego. Kilka wartości współczynników tarcia podaje tab. 6.2.

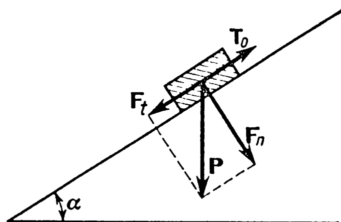
Tabela 6.2

Współczynniki tarcia statycznego μ_0 i kinetycznego μ

Rodzaj powierzchni ślizgających się	μ_0	μ
stal po lodzie	0,027	0,014
stal po stali	0,15	0,09–0,03
żeliwo po żeliwie	0,16	0,15
skóra po żelazie	0,28	0,27
dębina po dębinie (równolegle do włókien)	0,62	0,48
dębina po dębinie (prostopadle do włókien)	0,54	0,34
szkło po szkle		0,94
cegła po cegle	0,5–0,75	

Rozpatrzmy dla przykładu tarcie występujące podczas ruchu ciała po równi pochyłej, nachylonej pod kątem α do poziomu (rys. 6.16). Tym razem siłą $\mathbf{F}_n$ przyciskającą ciało do podstawy nie jest całkowity ciężar ciała $\mathbf{P}$ (tak byłoby w przypadku ruchu na płaszczyźnie poziomej), lecz tylko jego składowa prostopadła do powierzchni równi:

$$F_n = P\cos\alpha.$$



Rys. 6.16

W czasie zsuwania się po równi ciało o ciężarze $P = mg$ podlega zatem działaniu wypadkowej dwóch sił: siły zsuwającej $F_t = P\sin\alpha$ i hamującej siły tarcia $T = \mu P\cos\alpha$. Siła wypadkowa F wywołująca ruch równa się

$$F = P\sin\alpha - \mu P\cos\alpha$$

lub

$$F = mg(\sin\alpha - \mu\cos\alpha).$$

W czasie ruchu ciała po równi pochyłej do góry zarówno siła F_t, jak i siła tarcia T

są skierowane przeciwnie do prędkości. A zatem siła wypadkowa F, hamująca ruch, równa się w tym przypadku

$$F = mg(\sin\alpha + \mu\cos\alpha).$$

Stosując równię pochyłą o dającym się zmieniać kącie nachylenia można wyznaczać współczynnik tarcia statycznego. Jeżeli przy danym kącie nachylenia równi ciało samorzutnie nie zaczyna się zsuwać, to znaczy że siła tarcia T jest większa od siły zsuwającej F_t. Stopniowo zwiększając nachylenie można osiągnąć taki kąt α_t, przy którym ruch się rozpocznie. To świadczy o bardzo małej przewadze siły F_t nad siłą T. W przybliżeniu można zakładać, że przy tym kącie granicznym, zwanym *kątem tarcia*, zachodzi równość wspomnianych sił:

$$F_t = T_0,$$

gdzie T_0 oznacza siłę tarcia statycznego.

Ostatnie równanie można też napisać w postaci

$$P\sin\alpha_t = \mu_0 P\cos\alpha_t,$$

gdzie μ_0 oznacza współczynnik tarcia statycznego. Z dalszego przekształcenia wynika, że

$$\mu_0 = \operatorname{tg}\alpha_t.$$

Współczynnik tarcia statycznego równa się zatem tangensowi tego najmniejszego granicznego kąta nachylenia równi względem poziomu, przy którym ciało umieszczone na równi zaczyna samorzutnie się zsuwać.

Przejdźmy obecnie do tarcia tocznego. Łatwo można się przekonać, że tarcie występujące przy toczeniu się ciał jest mniejsze od tarcia poślizgowego, ale zawsze istnieje. Może warto też podkreślić, że na idealnie gładkiej powierzchni toczenie się w ogóle nie wystąpi. Na takiej powierzchni może odbywać się wyłącznie poślizg. Najczęściej w praktyce występują łącznie toczenie się i poślizg.

Tarcie występujące przy toczeniu jednego ciała po drugim zależy w dużym stopniu od właściwosci sprężystych toczącego się obiektu i podłoża. Zewnętrzna siła ciągnąca, podtrzymująca jednostajne toczenie się ciała, wykonuje równocześnie pracę odkształcania obiektu i podłoża. Doświadczalne prawo tarcia tocznego mówi o proporcjonalności momentu M siły ciągnącej względem linii styku toczącego się ciała z podłożem do siły przyciskającej ciało do podłoża. Innymi słowy

$$M = \mu_t F_n,$$

gdzie μ_t oznacza współczynnik tarcia tocznego.

Mimo pewnej analogii do prawa tarcia poślizgowego, widoczna jest wyraźna różnica w charakterze współczynników tarcia poślizgowego i tarcia tocznego. Pierwsza z wymienionych wielkości jest bezwymiarowa, druga — ma wymiar długości (dlaczego?).

Współczynniki tarcia tocznego przyjmują mniejsze wartości, gdy podłoże się mniej odkształca pod działaniem siły F_n. Widać to wyraźnie z przytoczonych niżej wartości μ_t (w metrach), odnoszących się do toczenia się kół po różnych podłożach:

szyny	0,00003
asfalt	0,0001

szosa I kl.	0,00016
bruk	0,0002–0,0003
piasek	0,0015–0,03

Tąrcie jest zjawiskiem czasem pożytecznym, lecz często też bardzo szkodliwym. Tarcie jako zjawisko pożyteczne umożliwia chodzenie, trzymanie przedmiotów, hamowanie pojazdów, walcowanie metali itp. Tarcie szkodliwe, występujące np. podczas ruchu maszyn, staramy się możliwie zmniejszać, gdyż obniża ono wydajność maszyn powodując zamianę części energii mechanicznej na rozpraszającą się bezużytecznie, energię cieplną i przyczynia się do szybszego niszczenia części lub nawet całości urządzeń. Zmniejszenie tarcia osiąga się przez usuwanie zanieczyszczeń z trących powierzchni, zamianę ruchu poślizgowego na toczenie (stosowanie np. łożysk kulkowych), rozdzielanie trących powierzchni warstwą cieczy (stosowanie smarów) itp. Ze względu na duże znaczenie zmniejszania tarcia dla współczesnej techniki badania jego mechanizmu są w wielu ośrodkach prowadzone bardzo intensywnie.

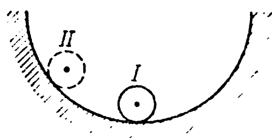
Pytania i zadania

1. Czym się zajmuje statyka?

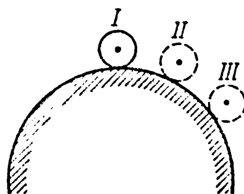
2. Czy stan ruchu lub spoczynku bryły sztywnej zmieni się po zastosowaniu dwóch dodatkowych sił równych co do wartości, lecz przeciwnie skierowanych?

3. Jaką siłą można zastąpić działanie na bryłę sztywną dwóch sił ukośnych leżących w jednej płaszczyźnie?

4. Podać dokładną charakterystykę siły wypadkowej w przypadku działania na ciało sztywne: a) dwóch sił równoległych zgodnych, b) dwóch sił równoległych, skierowanych przeciwnie, nierównych co do wartości.



Rys. 6.17



Rys. 6.18

5. Jakie są kinetyczne skutki działania na ciało sztywne kilku sił, których: a) wypadkowa siła i wypadkowy moment względem dowolnego punktu przestrzeni równają się zeru, b) wypadkowa siła nie równa się zeru, a wypadkowy moment względem dowolnego punktu przestrzeni równa się zeru, c) wypadkowa siła równa się zeru, a wypadkowy moment nie równa się zeru, d) wypadkowa siła i wypadkowy moment nie równają się zeru?

6. Podać dokładne wysłowienie warunku równowagi bryły sztywnej.

7. Czy przy ustalaniu warunków równowagi bryły sztywnej można uniknąć sumowania geometrycznego wszystkich sił działających na tę bryłę? Jaki układ równań skalarnych zastąpić może równanie wektorowe

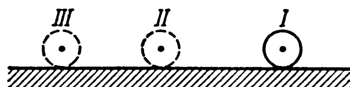
$$\sum_{i=1}^{n} \mathbf{F}_i = 0?$$

8. Czym należy się kierować przy wyborze — w zasadzie przecież dowolnego — punktu przestrzeni, względem którego obliczamy momenty kilku sił (znanych i nieznanych) podczas ustalania warunków równowagi bryły sztywnej?

9. Rysunki 6.17, 6.18 i 6.19 dotyczą trzech rodzajów równowagi w polu grawitacyjnym, a mianowicie równowagi trwałej, chwiejnej i obojętnej. Jak można określić te trzy rodzaje równowagi posługując się pojęciem: a) środka ciężkości, b) energii potencjalnej grawitacyjnej?

10. Jakie cechy ma siła tarcia poślizgowego?

11. Podać definicję współczynnika tarcia.



Rys. 6.19

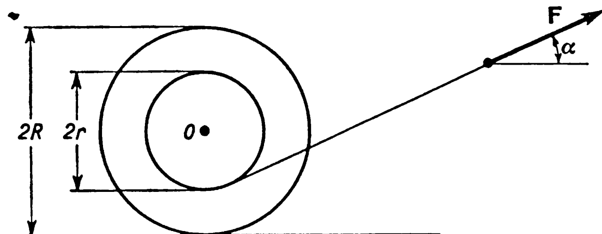
12. Omówić wyznaczanie współczynnika tarcia statycznego za pomocą równi pochyłej o zmienianym kącie nachylenia.

13. Co to jest tarcie toczne? Jakie czynniki decydują o wartości siły tarcia tocznego?

14. Wyjaśnić, dlaczego świeżo złapana ryba tak łatwo wysuwa się człowiekowi z ręki? Dlaczego zapałka ma chropowatą powierzchnię?

15. Jak wielkie są siły F_1 i F_2 wzajemnie do siebie prostopadłe, pozostające w stosunku $m:n$ i mające wypadkową równą F?

Odp. $F_1 = Fn/\sqrt{m^2+n^2}$, $F_2 = Fm/\sqrt{m^2+n^2}$.



Rys. 6.20

16. Między dwoma hakami odległymi od siebie o a rozpięto linkę o długości b. W środku linki zawieszono ciężar P. Zaniedbując ciężar linki obliczyć siłę, która ją rozciąga.

Odp. $Pb/2\sqrt{b^2-a^2}$.

17. Na dwóch kozłach położono deskę AD o długości L i ciężarze Q. Odcinek deski od jej końca A do punktu podparcia B ma długość a, odległość między punktami podparcia B i C równa się b. Między punktami podparcia w punkcie E odległym od punktu C o d stanął na desce człowiek ważący P kG. Jakie siły działają na kozły w obu punktach podparcia?

Odp. $F_1 = \dfrac{Pd+Q(a+b-L/2)}{b}$, $F_2 = Q+P-\dfrac{Pd+Q(a+b-L/2)}{b}$.

18. Drabina o długości L i o ciężarze Q jest oparta o gładką ścianę na wysokości h. Człowiek o ciężarze P stoi na stopniu znajdującym się na $\frac{1}{3}$ długości drabiny, licząc od podłogi. Zakładając, że środek ciężkości drabiny znajduje się w połowie jej długości, obliczyć siły, jakie drabina wywiera na ścianę i podłogę. (Uwaga: na podłogę działają dwie siły: styczna i normalna. Drabinę traktować jako urządzenie jednowymiarowe.)

Odp. $F_1 = F_{2t} = \dfrac{(3Q+2P)\sqrt{L^2-h^2}}{6h}$, $F_{2n} = P+Q$

19. O ścianę opiera się drabina, której środek ciężkości znajduje się w połowie jej długości. Współczynniki tarcia statycznego między drabiną a ścianą i podłogą wynoszą odpowiednio μ_1 i μ_2. Pod jakim najmniejszym kątem względem poziomu można ustawić drabinę nie wywołując jej upadku. Drabinę traktować jako urządzenie jednowymiarowe.

Odp. $\operatorname{tg}\alpha = \dfrac{1-\mu_1\mu_2}{2\mu_2}$.

20. Na poziomym stole leży szpulka nici o masie m, momencie bezwładności I i rozmiarach zaznaczonych na rys. 6.20. Z jakim przyspieszeniem liniowym przesuwa się środek masy szpulki, jeśli na koniec nitki działa siła **F** tworząca kąt α z poziomem. Jaki warunek musi spełniać kąt α, by szpulka toczyła się (bez poślizgu) w stronę działania siły **F**. Ile wynosi siła tarcia o stół?

Odp. $a = (R\cos\alpha - r)FR/(I+mR^2)$, $\cos\alpha > r/R$, $T = F\cos\alpha - \dfrac{mFR(R\cos\alpha - r)}{I+mR^2}$.

21. Sanki zsuwają się z góry o wysokości h i zatrzymują się na zaśnieżonym poziomym polu w odległości s liczonej od wierzchołka góry w kierunku poziomym. Wykazać, że współczynnik tarcia sanek o śnieg (jednakowy na zboczu góry i na polu) równa się stosunkowi h/s.

ROZDZIAŁ 7

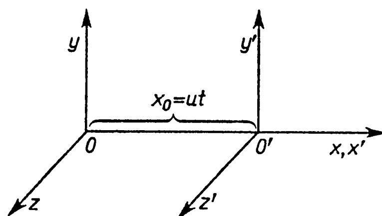
Elementy teorii względności

7.1. Rozważania wstępne

W naszych dotychczasowych rozważaniach braliśmy za podstawę mechanikę newtonowską, pamiętając jednak, że teoria stworzona przez Newtona, będąca niewątpliwie wielkim osiągnięciem jego geniuszu, jest obecnie traktowana jako przypadek szczególny teorii względności Einsteina, dotyczący ciał o prędkościach niewielkich w porównaniu z prędkością światła. Obecnie podkreślimy tło, na którym rozwinęła się tzw. szczególna teoria względności (1905 r.), omówimy pewne zmiany, które w stosunku do podstawowych pojęć i praw mechaniki wprowadził Einstein, oraz pewne wnioski wynikające z tej jego teorii. Ogólną teorią względności (1916 r.), będącą teorią grawitacji, nie będziemy się tutaj zajmowali.

7.2. Transformacje Galileusza

W paragrafie 2.6 zajmowaliśmy się zasadą względności mechaniki klasycznej. Wprowadziliśmy pojęcie inercjalnych układów odniesienia i stwierdziliśmy, że żaden z układów inercjalnych nie jest układem wyróżnionym: we wszystkich obowiązują te same prawa mechaniki, niezależnie od tego jak wielka jest prędkość danego układu,



Rys. 7.1

byleby była ona stała. Rozważmy raz jeszcze zależność między współrzędnymi opisującymi położenie punktu materialnego w układach X, Y, Z oraz X', Y', Z', przedstawionych na rys. 7.1, przy założeniach, że podczas ruchu układu primowanego X',

Y', Z': 1) osie Y' i Z' są stale równoległe do osi Y, Z, 2) początek układu O' przesuwa się wzdłuż osi X ze stałą prędkością u, 3) w chwili $t = 0$ początki obu układów pokrywają się ze sobą, 4) czas płynie jednakowo w obu układach, czyli $t = t_1$. W tych warunkach — w ujęciu niutonowskim — pewne zdarzenie, które w układzie X, Y, Z miało współrzędne czasowo-przestrzenne (x, y, z, t), będzie miało w układzie primowanym współrzędne:

$$x' = x - ut, \quad y' = y, \quad z' = z, \quad t' = t. \tag{7.1}$$

Równania te noszą nazwę *transformacji Galileusza.*

7.3. Transformacje Lorentza

Po przypomnieniu podstawowych równań, związanych z zasadą względności w mechanice klasycznej, powróćmy do teorii Einsteina i ustalmy tło, na jakim została ona stworzona. W związku z teorią względności, obowiązującą w mechanice klasycznej, nasunęło się pytanie, dlaczego w różnych układach inercjalnych stosują się bez zmian prawa mechaniki, a nie stosują się prawa elektrodynamiki. To samo doświadczenie elektryczne przedstawia się różnie w opisie obserwatora spoczywającego i obserwatora znajdującego się w ruchu. Odpowiedź mogła być różna:

1. Można było uważać, że zasada względności obowiązuje tylko w mechanice, a nie obowiązuje w elektrodynamice. Może w elektrodynamice istnieje tylko jeden wyróżniony układ odniesienia, w którym światło rozchodzi się z prędkością c.

2. Można też było kwestionować sformułowanie podstawowych praw elektrodynamiki. Może można by je przedstawić w takiej postaci, by zasada względności obowiązywała także w elektrodynamice.

3. Może należałoby zrewidować definicje pewnych podstawowych pojęć mechanicznych i ujęcie praw i dzięki wprowadzeniu pewnych modyfikacji osiągnąć ogólną stosowalność zasady względności zarówno w mechanice, jak i w elektryczności.

Z tych trzech dróg Einstein wybrał ostatnią — modyfikując pojęcia masy, pędu i energii, jak również długości i czasu, doszedł do tego, że zasada względności obejmuje zarówno prawa mechaniki, jak i elektrodynamiki. Trzeba zresztą podkreślić, że zarówno droga pierwsza, jak i druga nie znalazły potwierdzenia doświadczalnego (m. in. negatywny wynik słynnego doświadczenia Michelsona i Morleya). Wnioskiem z wykonanych doświadczeń było to, że *światło zawsze porusza się w próżni z prędkością c, niezależnie od ruchu źródła lub obserwatora.* Ten fakt fizyczny jest też podstawowym postulatem teorii Einsteina. Z niezależności prędkości światła od ruchu źródła lub obserwatora wynika, że pomiar prędkości światła dochodzącego od odległej gwiazdy da ten sam wynik, niezależnie od tego, czy jest wykonany przez obserwatora spoczywającego, czy też będącego w szybkim ruchu w kierunku gwiazdy. *Niezmienność wartości c wiąże się z tym, że w teorii względności obowiązują inne związki między współrzędnymi czasu i przestrzeni niż wyrażone równaniem* (7.1). Te nowe zależności noszą nazwę transformacji Lorentza.

Jak poprzednio, rozważamy układ X, Y, Z, który nazywać będziemy układem

laboratoryjnym, w którym obserwator A robi pomiary współrzędnych przestrzennych (x, y, z) oraz czasowych posługując się przy tym zegarem spoczywającym w jego układzie. Drugi obserwator B związany jest z układem (X', Y', Z'), poruszającym się ruchem jednostajnym z prędkością u w kierunku osi X z zachowaniem równoległości osi Y i Y' oraz Z i Z'. Każdemu zdarzeniu obserwator B przypisuje współrzędne przestrzenne x', y', z' oraz współrzędną czasową t' według pomiaru czasu dokonanego na zegarze takim samym jak w układzie laboratoryjnym, lecz poruszającym się wraz z układem (X', Y', Z'). Za początek rachuby czasu $t = t' = 0$ przyjmujemy czas, w którym oba układy pokrywają się. W tych warunkach, jak wykazał Lorentz (1904), obowiązują zależności

$$x = \frac{x'+ut'}{\sqrt{1-\frac{u^2}{c^2}}}, \quad x' = \frac{x-ut}{\sqrt{1-\frac{u^2}{c^2}}}, \tag{7.2}$$

$$y = y', \quad \text{lub} \quad y' = y,$$

$$z = z', \qquad z' = z,$$

$$t = \frac{t'+\left(\frac{u}{c^2}\right)x'}{\sqrt{1-\frac{u^2}{c^2}}}, \quad t' = \frac{t-\left(\frac{u}{c^2}\right)x}{\sqrt{1-\frac{u^2}{c^2}}}.$$

Są to właśnie transformacje Lorentza.

Zakładając słuszność transformacji Lorentza wykażemy, że prędkość światła w obu rozważanych układach jest jednakowa i równa c.

Niech źródło światła o współrzędnych $x = 0$, $y = 0$, $z = 0$ w chwili $t = 0$ wysyła sygnał w kierunku osi x. Wobec tego, że prędkość światła w układzie laboratoryjnym wynosi c, sygnał dotrze do punktu o współrzędnej x po czasie $t = x/c$. Odpowiednie współrzędne x' i t' w układzie obserwatora B (zgodnie z (7.2)) będą równe:

$$x' = \frac{x-u\frac{x}{c}}{\sqrt{1-\frac{u^2}{c^2}}} \quad \text{i} \quad t' = \frac{\frac{x}{c}-\left(\frac{u}{c^2}\right)x}{\sqrt{1-\frac{u^2}{c^2}}}.$$

Stąd prędkość światła v' w układzie B

$$v' = \frac{x'}{t'} = \frac{x-\frac{ux}{c}}{\sqrt{1-\frac{u^2}{c^2}}}\,\frac{\sqrt{1-\frac{u^2}{c^2}}}{\frac{x}{c}-\left(\frac{u}{c^2}\right)x} = \frac{x-\frac{ux}{c}}{\frac{1}{c}\left(x-\frac{ux}{c}\right)} = c.$$

Tak więc postulat Einsteina dotyczący niezależności prędkości światła c od ruchu obserwatora jest spełniony, gdy opieramy się na transformacjach Lorentza. Ogólnie

biorąc, można powiedzieć, że *sens matematyczny teorii względności sprowadza się do niezmienniczości równań opisujących prawa fizyczne względem transformacji Lorentza.*

W dalszych punktach omówimy ważne konsekwencje wypływające z transformacji Lorentza.

7.4. Relatywistyczne dodawanie prędkości

Jedną z konsekwencji transformacji Lorentza jest fakt, że *należy odróżniać newtonowskie dodawanie prędkości od relatywistycznego.* Z poprzednich rozważań (§ 2.6) wiemy, że prędkość $\mathbf{v}$ ciała względem układu laboratoryjnego w ujęciu Newtona jest sumą geometryczną prędkości $\mathbf{v}'$ tego ciała względem układu primowanego i prędkości unoszenia $\mathbf{u}$ tego układu. A zatem przy takim wyborze układów, jaki przyjęliśmy poprzednio, newtonowskie dodawanie prędkości sprowadzi się do zależności

$$v = v' + u. \tag{7.3}$$

Spróbujmy znaleźć relatywistyczny odpowiednik tego równania. Niech nasi obserwatorzy A i B obserwują ruch punktu materialnego w swoich układach. Zdaniem pierwszego, $v = dx/dt$, zdaniem drugiego, $v' = dx'/dt'$. Ale x' i t' można wyrazić w zależności od x i t, a zatem

$$v' = \frac{dx'}{dt'} = \frac{\dfrac{d(x-ut)}{\sqrt{1-\dfrac{u^2}{c^2}}}}{\dfrac{d\left(t-\dfrac{u}{c^2}x\right)}{\sqrt{1-\dfrac{u^2}{c^2}}}} = \frac{dx-udt}{dt-\dfrac{u}{c^2}dx}.$$

Dzieląc licznik i mianownik przez dt, znajdujemy

$$v' = \frac{v-u}{1-\dfrac{u}{c^2}v}, \tag{7.4}$$

skąd

$$v = \frac{v'+u}{1+\dfrac{u}{c^2}v'}. \tag{7.5}$$

Zamiast zależności newtonowskiej (7.3) otrzymaliśmy zależność relatywistyczną (7.5) przechodzącą w (7.3) w przypadku, gdy mamy do czynienia z prędkościami bardzo małymi w porównaniu z prędkością światła c.

Ze wzoru określającego relatywistyczne dodawanie prędkości wynika bardzo ważny wniosek: prędkość obiektu nigdy nie przekracza prędkości światła c. Zbadajmy to na przykładzie. Niech obserwator B razem z układem primowanym porusza się względem układu laboratoryjnego z prędkością $u = 0{,}9c$. Według jego pomiaru jakiś porusza-

jący się przedmiot ma prędkość v' równą np. także $0{,}9c$. Prędkość v tego przedmiotu względem układu laboratoryjnego nie równa się $1{,}8c$, lecz wynosi zgodnie z (7.5)

$$v = \frac{0{,}9c+0{,}9c}{1+\frac{(0{,}9c)^2}{c^2}} = \frac{1{,}8c}{1{,}81} \cong 0{,}994c.$$

Interesujący też jest przykład dotyczący relatywistycznego obliczania prędkości, jeśli jedna z nich, a mianowicie v równa się prędkości fotonu, czyli prędkości światła c, a druga stanowi prędkość u obserwatora względem np. źródła fotonów. Wtedy prędkość v', jaką obserwator przypisuje fotonom, wynosi zgodnie z (7.4)

$$v' = \frac{c-u}{1-\frac{cu}{c^2}} = c.$$

Innymi słowy, każdy obserwator — niezależnie od własnej prędkości u — stwierdzi, że światło porusza się z prędkością c.

7.5. Zmiana masy ciała w zależności od prędkości

Podstawowym wnioskiem wynikającym z zasad dynamiki Newtona jest zasada zachowania pędu w układach odosobnionych. W klasycznym ujęciu tej zasady przyjmuje się, że masy działających na siebie ciał, wchodzących w skład układu, są stałe.

Tabela 7.1

Zmienność masy ciał w zależności od prędkości

v/c	m/m_0
0	1,00000
0,01	1,00005
0,02	1,00020
0,05	1,00125
0,1	1,00504
0,2	1,02062
0,5	1,1547
0,6	1,2500
0,7	1,4002
0,8	1,6666
0,9	2,2941
0,99	7,0888

Gdyby zachować to założenie, to w ujęciu relatywistycznym pęd układu ulegałby zmianie. Dla utrzymania słuszności zasady zachowania pędu w układach odosobnionych należy zmienić definicję pędu (lub masy), a mianowicie należy przyjąć, że pęd

$$p = mv = \frac{m_0 v}{\sqrt{1-\frac{v^2}{c^2}}}, \tag{7.6}$$

czyli że masa m ciała będącego w ruchu równa się

$$m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}, \tag{7.7}$$

gdzie m_0 oznacza masę ciała w spoczynku, v — prędkość ciała. Dla prędkości v małych w porównaniu z prędkością światła c można przyjmować, że $m = m_0$.

Wartości występujące w tab. (7.1) wskazują, że efekty relatywistyczne można zaniedbywać, gdy prędkości v nie przekraczają $0{,}1c$.

Jest rzeczą jasną, że z równania (7.7) można obliczać prędkość ruchu ciała, jeśli znany jest stosunek m_0/m.

7.6. Związek masy ciała z energią

Przy klasycznym ujęciu energii kinetycznej ciała mówiliśmy, że E_k równa się pracy potrzebnej do nadania ciału pierwotnie spoczywającemu ($v_0 = 0$) prędkości końcowej v. A zatem

$$E_k = \int F dx.$$

Ale siłę można przedstawić jako pochodną pędu względem czasu, a zatem zgodnie z (7.6)

$$E_k = \int_0^v \frac{d}{dt} \frac{m_0 v}{\sqrt{1-\frac{v^2}{c^2}}} dx = m_0 \int_0^v v d \frac{v}{\sqrt{1-\frac{v^2}{c^2}}}.$$

Całkując to wyrażenie przez części otrzymamy

$$E_k = mc^2 - m_0 c^2 \tag{7.8}$$

lub

$$E_k = \Delta mc^2. \tag{7.9}$$

Równanie powyższe zastępuje klasyczne wyrażenie na energię kinetyczną $E_k = \frac{1}{2} mv^2$ w przypadku dużych prędkości v.

Powróćmy raz jeszcze do równania (7.8). Można je napisać w postaci

$$mc^2 = E_k + m_0 c^2.$$

Przyjmując za Einsteinem, że iloczyn masy relatywistycznej i kwadratu prędkości światła przedstawia całkowitą energię E ciała, otrzymujemy

$$E = mc^2, \tag{7.10}$$

$$E = E_k + m_0 c^2.$$

To znaczy, że *całkowita energia E ciała składa się z energii kinetycznej E_k i energii spoczynkowej m_0c^2*, zwanej przez Einsteina energią wewnętrzną spoczynkową.

Zgodnie z przewidywaniami Einsteina możliwe jest (przy zachowaniu pewnych warunków) przekształcanie się energii spoczynkowej m_0c^2 w inne rodzaje energii. Do tych spraw powrócimy jeszcze w rozdz. 31.

Z równań (7.8) i (7.9) nie widać bezpośrednio zależności E_k od prędkości v ruchu ciała, jak również tego, że dla $v \ll c$ wzory te przechodzą w klasyczny wzór określający energię kinetyczną. Zbadajmy te sprawy dokładniej. Z (7.8) otrzymujemy

$$E_k = \frac{m_0c^2}{\sqrt{1-\frac{v^2}{c^2}}} - m_0c^2,$$

$$E_k = m_0c^2\left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}-1\right) = m_0c^2\left[\left(1-\frac{v^2}{c^2}\right)^{-1/2}-1\right].$$

Z rozwinięcia dwumianu w szereg wynika, że

$$\left(1-\frac{v^2}{c^2}\right)^{-1/2} = 1+\tfrac{1}{2}\frac{v^2}{c^2}+\ldots,$$

gdyż $(1+a)^n = 1+na+\frac{n(n-1)}{1\cdot 2}a^2+\ldots$

Przy założeniu, że $v \ll c$, dalsze wyrazy rozwinięcia naszego wyrażenia zaniedbujemy. A zatem

$$E_k = m_0c^2\left(1+\tfrac{1}{2}\frac{v^2}{c^2}-1\right) = \tfrac{1}{2}m_0v^2.$$

7.7. Skrócenie Lorentza i paradoks bliźniąt

Na zakończenie tego krótkiego omówienia elementów szczególnej teorii względności zwrócimy jeszcze uwagę na dwa fakty. Pierwszy dotyczy pomiaru długości w układzie laboratoryjnym i w układzie poruszającym się względem niego ze stałą prędkością u, drugi zaś — pomiaru czasu dokonywanego za pomocą identycznych zegarów w obu wspomnianych układach.

Niech przedmiotem mierzonym będzie np. taśma miernicza o długości L' metrów w układzie poruszającym się, umieszczona wzdłuż osi x tak, że jej początek przypada w początku tego układu. A zatem współrzędna początku taśmy równa się $x'_0 = 0$, a końca taśmy $x'_k = L'$.

Zgodnie z transformacją Lorentza możemy napisać, że

$$x'_0 = \frac{x_0-ut}{\sqrt{1-\frac{u^2}{c^2}}} = 0, \tag{7.11}$$

$$x'_k = \frac{x_k - ut}{\sqrt{1-\frac{u^2}{c^2}}} = L'. \tag{7.12}$$

Z (7.11) wynika, że

$$x_0 = ut,$$

co po podstawieniu do (7.12) daje

$$x_k - x_0 = L'\sqrt{1-\frac{u^2}{c^2}}.$$

Ale $x_k - x_0$ jest długoscią naszej taśmy mierniczej według pomiaru obserwatora z układu laboratoryjnego. W stosunku do L' jest ona zmniejszona $\sqrt{1-u^2/c^2}$ razy.

Warto podkreślić, że to *skrócenie długości* (zwane często *skróceniem Lorentza*) *dotyczy wymiaru ciała w kierunku równoległym do kierunku prędkości u układu*. W kierunku prostopadłym do prędkości skrócenie nie występuje. Jeżeli np. w układzie primowanym porusza się kula o promieniu r, to w układzie laboratoryjnym stwierdzony zostanie ruch elipsoidy obrotowej, której oś w kierunku ruchu wynosi $r\sqrt{1-u^2/c^2}$, a pozostałe dwie osie mają długość r.

Przejdźmy teraz do zagadnień związanych z pomiarem czasu w rozważanych układach. Załóżmy, że dwa zdarzenia odbywają się w układzie primowanym w tym samym punkcie (x', y', z'), lecz w różnych czasach t'_1 i t'_2. Obserwator z układu laboratoryjnego określi te zdarzenia jako zachodzące w różnych punktach przestrzeni i w różnych czasach. Współrzędne przestrzenne w układzie laboratoryjnym będą (x_1, y, z) oraz (x_2, y, z), współrzędne czasowe t_1 i t_2. Łatwo sprawdzić, że spełnione są zależności

$$x_2 - x_1 = u(t_2 - t_1),$$

jak również (według transformacji Lorentza)

$$t_2 - t_1 = \frac{t'_2 - t'_1}{\sqrt{1-\frac{u^2}{c^2}}}.$$

A zatem sekwencja obu zdarzeń w czasie jest w obu układach jednakowa, lecz odstępy czasowe między zdarzeniami są nierówne

$$\Delta t \neq \Delta t'.$$

Odstęp czasu Δt w układzie laboratoryjnym jest dłuższy od $\Delta t'$ w układzie primowanym. Innymi słowy, wydaje się, że zegar poruszający się z układem primowanym „chodzi" wolniej niż zegar spoczywający w układzie laboratoryjnym. Zmiana czasu zachodzi w stosunku $\sqrt{1-u^2/c^2}:1$.

Zastanówmy się wobec tego, co by się stało, gdyby z dwojga bliźniąt z układu laboratoryjnego jeden podjął decyzję odbycia podróży kosmicznej z dużą prędkością. Po powrocie na Ziemię stwierdziłby on, że jego brat postarzał się bardziej niż on sam. Byłoby to skutkiem „rozszerzenia" czasu w jego pojeździe kosmicznym.

Z zasady względności wynika, że opis zdarzenia we wszystkich układach inercjalnych powinien być taki sam. Szybsze starzenie się bliźniaka z układu laboratoryjnego

wydaje się temu przeczyć. Paradoks jest jednak tylko pozorny, gdyż w rozważanym przypadku układ poruszającego się bliźniaka nie jest układem inercjalnym: zarówno na początku swego ruchu, jak i przy zmianie kierunku ruchu układ porusza się z przyspieszeniem. Bliźniak poruszający się nie ogląda swego brata na początku i przy końcu swej wycieczki z tego samego inercjalnego układu, a zatem opis zdarzeń w ujęciu obu bliźniaków musi być różny.

7.8. Podsumowanie wiadomości

Kończąc ten krótki przegląd podstawowych wiadomości z teorii względności zestawmy najważniejsze punkty.

1. *Celem*, jaki przyświecał Einsteinowi przy opracowywaniu szczególnej teorii względności, było *ustalenie takich zależności czasowo-przestrzennych, charakteryzujących przebieg zjawisk w układach inercjalnych, by zachowana była niezmienniczość wszelkich praw fizycznych w tych układach.*

2. *Drogą* do uzyskania *celu*, wymienionego wyżej, *było zastąpienie transformacji Galileusza transformacjami Lorentza* (7.2) oraz *przyjęcie postulatu opartego na fakcie doświadczalnym, że prędkość światła w próżni we wszystkich układach inercjalnych jest taka sama i równa c.*

3. Z transformacji Lorentza wynika szereg bezpośrednich *konsekwencji*, jak np.:

a) *relatywistyczne dodawanie prędkości* — zamiast wzoru mechaniki klasycznej (7.3) określającego dodawanie prędkości obowiązuje wzór (7.5). Konsekwencją tego wzoru jest niemożliwość wystąpienia ruchu ciała z prędkością większą od prędkości światła c.

b) *skrócenie długości* — długość L' zmierzona w układzie primowanym oceniana jest w układzie laboratoryjnym jako równa

$$L = L'\sqrt{1-\frac{u^2}{c^2}}.$$

c) *wydłużenie czasu* — odstęp czasu $\Delta t'$ zmierzony na zegarze w układzie primowanym oceniany jest w układzie laboratoryjnym (według pomiaru na zegarze w tym układzie) jako

$$\Delta t = \frac{\Delta t'}{\sqrt{1-\frac{u^2}{c^2}}}.$$

4. *Dla zachowania niezmienniczości praw dynamiki konieczne było przyjęcie zmienności masy w zależności od prędkości według wzoru*

$$m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}.$$

W związku z tym *pęd w mechanice relatywistycznej wyraża się wzorem*

$$p = \frac{m_0 v}{\sqrt{1-\frac{v^2}{c^2}}}.$$

5. Jednym z ważniejszych wyników szczególnej teorii względności jest odkrycie przez Einsteina *równoważności masy i energii*:

$$E = mc^2.$$

Stąd wniosek, że energia doprowadzona do ciała powiększa jego masę i odwrotnie, energia wydana przez ciało zmniejsza jego masę. Tak np. Słońce wypromieniowując energię zmniejsza stale swą masę o wiele milionów ton rocznie.

Pytania i zadania

1. Jakie wielkości fizyczne uległy modyfikacji w teorii Einsteina?
2. Na czym polegają podstawowe różnice między transformacjami Galileusza i transformacjami Lorentza?
3. Czy przyspieszenie określonego ciała uzyskane pod działaniem określonej siły jest zależne od ruchu ciała?
4. Czy klasyczne i relatywistyczne dodawanie prędkości prowadzi do tych samych wyników?
5. Czy długość i czas mają znaczenie absolutne, czy też zależą od ruchu przyrządów pomiarowych?
6. W jakich warunkach odstępstwa od klasycznej mechaniki newtonowskiej są duże? Kiedy należy stosować mechanikę relatywistyczną?
7. Jaki jest związek masy ciała z energią?
8. Z czego się składa całkowita energia E badanego ciała?
9. Na czym polega skrócenie Lorentza?
10. Czy w związku ze skróceniem Lorentza występuje zmiana kształtu przedmiotu obserwowanego w układzie primowanym i układzie laboratoryjnym?
11. Wyjaśnić podstawy paradoksu bliźniąt.
12. Na jakiej podstawie współczynnik $\beta = \sqrt{1-v^2/c^2}$ występujący w bardzo wielu równaniach relatywistycznych można zastępować wyrażeniem $1-\frac{1}{2}(v^2/c^2)$, jeżeli $v \ll c$?
13. Wyrazić jednostkę masy atomowej (tj. $\frac{1}{12}$ masy spoczynkowej atomu węgla $^{12}C = 1{,}66 \cdot 10^{-27}$ kg) w jednostkach energii.

Odp. $1{,}49 \cdot 10^{-10}$ J.

14. Obliczyć prędkość, z jaką powinna by się poruszać jednostka masy atomowej, aby jej energia kinetyczna odpowiadała podwojonej masie spoczynkowej.

Odp. $2{,}83 \cdot 10^8$ m/s.

15. Jaki otrzymalibyśmy wynik rozwiązując poprzednie zadanie nierelatywistycznie?

Odp. $2c$.

16. W jakim stosunku do energii spoczynkowej pozostaje praca, którą należy wykonać, aby prędkość cząstki zwiększyć od $0{,}90c$ do $0{,}99c$.

Odp. $5{,}81 m_0c^2$.

17. Obliczyć długość podziałki metrowej, która się porusza względem obserwatora z prędkością $0{,}90c$. Rozważyć przypadki, gdy podziałka jest zorientowana: a) równolegle, b) prostopadle do kierunku swej prędkości.

Odp. 18,5 cm, 100 cm.

18. Obliczyć energię wyzwoloną w czasie połączenia się pary elektron–pozyton.

Odp. $1{,}632 \cdot 10^{-13}$ J.

19. Znaleźć wynik składania dwóch prędkości skierowanych w tę samą stronę, jednakowych co do wartości i równych $\frac{2}{3}c$.

Odp. $\frac{12}{13}c$.

ROZDZIAŁ 8

Właściwości sprężyste ciał

8.1. Ogólna charakterystyka ciał odkształcalnych

W dotychczasowych naszych rozważaniach ograniczaliśmy się do badania działania sił na punkty materialne i bryły sztywne. W ten sposób abstrahowaliśmy od takich skutków działania sił, jak zmiana kształtu lub objętości. W rozdziale niniejszym i następnym zestawimy podstawowe wiadomości dotyczące sprężystych właściwości ciał, które w zależności od budowy wewnętrznej dzielimy na jednorodne i niejednorodne oraz izotropowe i anizotropowe. *Ciała jednorodne* są to takie ciała, w których każdy element objętości ma jednakowe właściwości fizykochemiczne (m.in. właściwości mechaniczne, elektryczne, optyczne). Ciała izotropowe są to takie ciała, których właściwości fizykochemiczne są jednakowe we wszystkich kierunkach pomyślanych w ciele. *W ciałach anizotropowych* właściwości te zależą od kierunków wybranych w ciele. Do ciał jednorodnych anizotropowych należą m.in. czyste ciała krystaliczne z wyjątkiem kryształów układu regularnego. Przykładem ciała niejednorodnego anizotropowego może być drewno.

Zjawisko odkształceń sprężystych ciał niejednorodnych i anizotropowych jest zjawiskiem złożonym, zależnym — jak wykazują badania — od wielu parametrów charakteryzujących właściwości sprężyste ciał. Mimo że ciała rzeczywiste są raczej niejednorodne i często anizotropowe, w dalszych rozważaniach ograniczymy się do uwag ogólnych dotyczących sprężystości ciał jednorodnych i izotropowych. Do tej kategorii można w przybliżeniu zaliczać metale, szkło, wiele mas plastycznych itp.

Działaniu sił zewnętrznych, powodujących zmiany objętości lub kształtu badanego ciała, towarzyszy *zakłócenie rozkładu sił międzycząsteczkowych*, objawiające się w pojawieniu się sił wewnętrznych, zwanych *siłami sprężystości*. W przypadku małych sił występuje równoważenie się sił zewnętrznych i sił wewnętrznych. Po usunięciu takich sił zewnętrznych ciało wraca do pierwotnego kształtu i objętości. W tych warunkach mówimy o *odkształceniach sprężystych* (lub *doskonale sprężystych*).

Przy wzroście siły zewnętrznej aż do pewnej wartości granicznej rośnie też stopniowo odkształcenie badanego ciała zachowując swój charakter sprężysty: wywołane odkształcenia należą do *zakresu* (*obszaru*) *właściwości sprężystych ciała*. Po przekroczeniu granicznej wartości siły zewnętrznej (zależnej od rodzaju badanego ciała) powstają *od-*

kształcenia niesprężyste. Po usunięciu sił zewnętrznych siły wewnętrzne nie sprowadzają ciała do pierwotnego kształtu i objętości. Pozostaje w ciele odkształcenie trwałe. W tych warunkach mówimy o odkształceniach należących do *zakresu* (*obszaru*) *właściwości plastycznych ciała*.

Przy omawianiu właściwości sprężystych ciał zamiast pojęć siły zewnętrznej i wewnętrznej stosuje się najczęściej pojęcie ciśnienia (zewnętrznego) i naprężenia (wewnętrznego). Wielkości te kojarzymy z działaniem siły (odpowiednio zewnętrznej lub wewnętrznej) na jednostkę powierzchni. Innymi słowy, *ciśnienie* (*naprężenie*) p w przypadku równomiernego rozkładu siły F na powierzchni S wynosi

$$p = \frac{F}{S} \tag{8.1}$$

lub przy siłach o rozkładzie nierównomiernym

$$p = \lim_{\Delta S \to 0} \frac{\Delta F}{\Delta S} = \frac{dF}{dS}, \tag{8.2}$$

gdzie ΔF oznacza siłę działającą na mały element powierzchni ΔS. Naprężenia i ciśnienia wyrażamy w tych samych jednostkach. *W układzie SI jednostką naprężenia i ciśnienia* jest niuton na metr kwadratowy (N/m^2), zwany też paskalem*.

8.2. Odkształcenia sprężyste. Prawo Hooke'a

Zanim przejdziemy do sformułowania podstawowego prawa dotyczącego właściwości sprężystych ciał, scharakteryzujemy kilka rodzajów odkształceń sprężystych.

Odkształcenia mogą mieć różny charakter. Można tu wymienić wydłużenia (skrócenia), rozszerzenia objętościowe (skurczenia), zgięcia, skręcenia itp. Omówimy bardziej szczegółowo trzy rodzaje odkształceń, a mianowicie odkształcenia związane ze zmianą długości (tzw. wydłużenia), ze zmianą objętości (tzw. odkształcenia objętościowe) i ze zmianą kształtu (tzw. odkształcenia postaciowe).

Za miarę odkształcenia związanego ze zmianą długości przyjmujemy względny przyrost długości ε, zdefiniowany jako stosunek przyrostu długości Δl (np. pręta) do długości pierwotnej l_0 (rys. 8.1):

$$\varepsilon = \frac{\Delta l}{l_0}.$$

Idealne odkształcenie objętościowe występuje wtedy, gdy kształt ciała zostaje zachowany, natomiast gęstość ulega zmianie (np. ciało kuliste pozostaje ciałem kulistym, ale ma już mniejszą lub większą objętość) (rys. 8.2). *Za miarę odkształcenia objętościowego* przyjmujemy względny przyrost objętości θ zdefiniowany jako stosunek przyrostu objętości ΔV do objętości pierwotnej V_0:

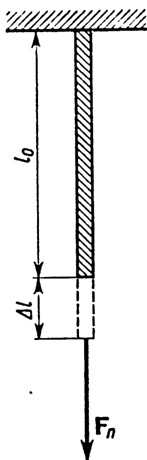
$$\theta = \frac{\Delta V}{V_0}.$$

---

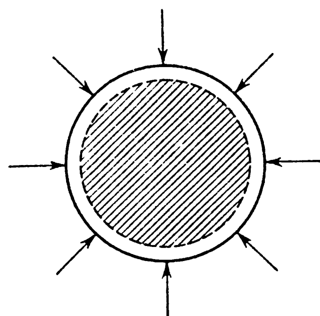
* Liczba dotychczas stosowanych jednostek ciśnienia była bardzo duża. Wymienimy kilka z nich: $dyna/cm^2$, bar, G/cm^2, atmosfera fizyczna i techniczna, milimetr słupa rtęci itp. (por. tab. U.1).

Z idealnym odkształceniem postaciowym mamy do czynienia wtedy, gdy gęstość dowolnie małego elementu ciała pozostaje niezmienna, a zmianie ulega kształt. Poglądowo takie odkształcenie przedstawione jest na rys. 8.3. Widzimy tam kostkę sześcienną, przymocowaną nieruchomo do podstawy. Na powierzchnię $BDHE$ działa stycznie siła $\mathbf{F}_t$, wywołując przesunięcie warstwy górnej ciała względem niżej położonych. Towarzyszy temu przesunięcie ścian bocznych o kąt γ względem pierwotnego położenia. *Za miarę tego odkształcenia* (zwanego też ścinaniem lub *skręceniem prostym*) przyjmiemy

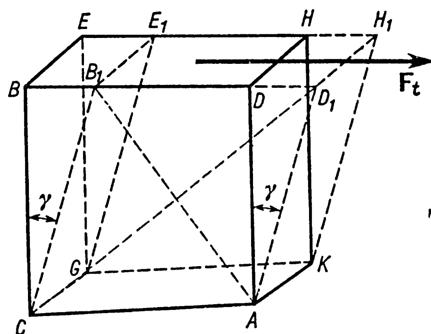
$$\operatorname{tg}\gamma = \frac{BB_1}{BC}.$$



Rys. 8.1



Rys. 8.2



Rys. 8.3

Przy małych odkształceniach postaciowych można korzystać z przybliżonej równości

$$\operatorname{tg}\gamma \approx \gamma.$$

Łatwo sprawdzić, że jeśli mamy do czynienia z idealnym odkształceniem postaciowym, to następuje wydłużenie w kierunku przekątnej CD_1 oraz skrócenie w kierunku przekątnej AB_1.

Warto podkreślić, że aczkolwiek odkształcenia sprężyste występujące w przyrodzie mają na ogół charakter złożony, to jednak dają się one sprowadzić do odkształceń objętościowych i postaciowych. Z tym się wiąże znaczenie obu tych idealnych rodzajów odkształceń. Odpowiednio do tych dwóch rodzajów odkształceń rozróżniamy też dwa rodzaje sprężystości: sprężystość objętościową i postaciową.

Przejdziemy obecnie do omówienia podstawowego prawa rządzącego sprężystością ciał. Jest to *prawo doświadczalne* podane przez *Hooke'a* już w roku 1675 i *obowiązujące dla małych odkształceń*. Prawo to mówi, że *stosunek naprężenia do związanego z nim odkształcenia jest wielkością stałą dla danego materiału.* Stosunek ten nazywamy *modułem sprężystości.*

Prawo Hooke'a przy wydłużeniach wyraża się wzorem

$$\sigma = E\frac{\Delta l}{l_0} = E\varepsilon, \tag{8.3}$$

gdzie σ oznacza ciśnienie (naprężenie) normalne, wynikające z działania sił prostopadłych do powierzchni, E — moduł sprężystości przy wydłużeniu, zwany *modułem Younga.*

Znajomość modułu Younga dla danego materiału pozwala obliczyć przyrost długości prętów lub drutów, wywołany działaniem siły normalnej F_n (rys. 8.1). Opierając się na wzorach (8.1) i (8.3) można napisać

$$\frac{F_n}{S} = E\frac{\Delta l}{l_0}, \tag{8.4}$$

gdzie S oznacza przekrój poprzeczny badanego pręta lub drutu. Z (8.4) otrzymujemy

$$\Delta l = \frac{l_0 F_n}{ES}.$$

W odniesieniu do odkształceń objętościowych prawo Hooke'a możemy napisać w postaci

$$\sigma = -K\frac{\Delta V}{V_0} = -K\theta, \tag{8.5}$$

gdzie σ oznacza, jak poprzednio, ciśnienie (naprężenie) normalne do powierzchni, K — moduł sprężystości objętościowej. Znak minus występujący w równaniu przypomina, że dodatnim przyrostom ciśnień (naprężeń) odpowiadają ujemne przyrosty objętości: im większe zastosujemy ciśnienie ściskające, tym bardziej ciało się skurczy.

W przypadku odkształceń postaciowych prawo Hooke'a wyraża się wzorem

$$\tau = G\operatorname{tg}\gamma, \tag{8.6}$$

gdzie τ oznacza ciśnienie styczne do powierzchni, G — moduł sprężystości postaciowej.

Ze wzorów (8.3), (8.5) i (8.6) wynika, że *moduły sprężystości wyrażają się w jednostkach ciśnienia* (wielkości ε, θ i $\operatorname{tg}\gamma$ są niemianowane).

Pozostawiając bliższe omówienie właściwości sprężystych ciał ciekłych i gazowych

do następnego rozdziału ograniczymy się tutaj do rozszerzenia wiadomości o właściwościach sprężystych ciał stałych*.

Wartości liczbowe modułów sprężystości ciał stałych (por. tab. 8.1) są bardzo duże ze względu na małą ich odkształcalność. Aby to lepiej zrozumieć, załóżmy, że zmiana długości Δl równa się pierwotnej długości l_0. Wówczas zgodnie ze wzorem (8.3) zastosowane ciśnienie równa się modułowi Younga. Innymi słowy, *moduł Younga* wyraża wartość ciśnienia, jakie należałoby zastosować, aby długość ciała zwiększyć o 100%, pozostając w zakresie proporcjonalności ciśnienia do związanego z nim odkształcenia (czyli w zakresie stosowalności prawa Hooke'a). Wywołanie takiego odkształcenia wymagałoby bardzo wielkiego ciśnienia — stąd wynika duża wartość modułu sprężystości na wydłużenie. W praktyce takich odkształceń na ogół nie można uzyskać, gdyż już wcześniej tracą one charakter sprężysty. Do takiego eksperymentu nadawałby się kauczuk.

Tabela 8.1

Wielkości charakteryzujące właściwości sprężyste różnych materiałów (temp. 15°C)

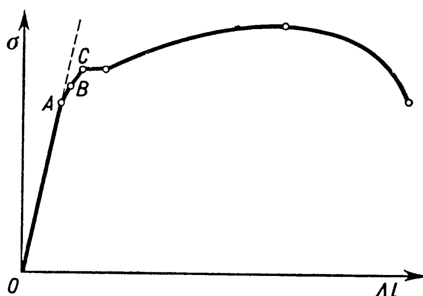
Rodzaj materiału	Moduły			Granice	
	Younga E	sprężystości postaciowej G	sprężystości objętościowej K	sprężystości	wytrzymałości
	10^{11} N/m²			10^{7} N/m²	
wolfram	3,62	1,35	3,3		420–700
stal	2,19	0,83	1,7	17–28	38–190
nikiel	2,03	0,79	2,0	14–21	35–49
miedź	0,8–1,3	0,4–0,49	1,43		40–45
brąz	1,05–1,2	3,2		11–23	2,5–62
mosiądz	1,05	0,43	1,0	39–47	50–61
glin	0,6–0,7	0,23–0,27	0,74	10–13	17–27
ołów	0,15	0,065			1,7–2,2

Właściwości sprężyste ciała podczas rozciągania można zilustrować np. za pomocą rys. 8.4. Rysunek odnosi się do stali tzw. niskowęglowej, wyżarzonej. Na osi odciętych odłożone są bezwzględne przyrosty długości Δl, na osi rzędnych — naprężenia normalne σ. Dla naprężeń rosnących stopniowo od zera otrzymuje się na wykresie odcinek linii prostej OA, odpowiadający prostej proporcjonalności między naprężeniem i wydłużeniem. Odcinek ten odpowiada zakresowi dokładnej stosowalności prawa Hooke'a. Wartość naprężenia będącą rzędną punktu A nazywamy granicą proporcjonalności. Innymi słowy, *granicą proporcjonalności* jest wielkość mianowana, wyrażająca maksymalne

* Warto zwrócić uwagę na różnice znaczenia pojęcia „ciało stałe" w mowie potocznej i w terminologii fizycznej. W mowie potocznej ciało stałe oznacza zarówno ciało krystaliczne, jak i bezpostaciowe. *W fizyce do ciał stałych zaliczane są tylko ciała krystaliczne. Ciała bezpostaciowe traktowane są jako przechłodzone ciecze.* Bliższe wiadomości na ten temat podane są w rozdz. 16.

naprężenie, przy którym jeszcze stosuje się prawo Hooke'a. Naprężenia o wartościach od zera do granicy proporcjonalności wywołują odkształcenia sprężyste, należące do tzw. zakresu (obszaru) proporcjonalności (zakresu stosowalności prawa Hooke'a).

Stosowanie ciśnień większych od granicy proporcjonalności odpowiada gwałtowniejszemu wzrostowi przyrostów długości, niż to miało miejsce w obszarze proporcjonalności: na wykresie począwszy od punktu A nie mamy już dalszego prostoliniowego przebiegu. Ciało jednak nadal, aż do wystąpienia naprężenia odpowiadającego punktowi B wykresu, zachowuje swe właściwości sprężyste, tzn. po usunięciu sił zewnętrznych, dzięki działaniu sił wewnętrznych, wraca do pierwotnej długości.



Rys. 8.4

Wartość naprężenia odpowiadającą punktowi B wykresu nazywamy *granicą sprężystości*. Jest to zatem znowu wielkość mianowana, wyrażająca maksymalne naprężenie, przy którym jeszcze ciało zachowuje się jak ciało sprężyste. Odkształcenie towarzyszące tej wartości naprężenia zamyka tzw. zakres (obszar) sprężystości. Zastosowanie ciśnień zewnętrznych większych od granicy sprężystości powoduje wystąpienie odkształceń trwałych. Przy dalszym wzroście ciśnienia ciało nabiera właściwości plastycznych, jest skłonne do „płynięcia", tzn. wzrost długości występuje bez wzrostu naprężenia. Mówimy wtedy o *obszarze plastyczności*. Obszar ten jest wykorzystywany przy obróbce ciał dla stosunkowo łatwego nadawania im żądanych kształtów (np. przy tłoczeniu). Płynięcie materiału jest jednak tylko cechą przejściową. Do wywołania dalszego wzrostu wydłużenia potrzebny jest wzrost naprężenia aż do chwili, gdy na pręcie pojawia się przewężenie — „szyjka". Następuje dalszy wzrost długości i wreszcie dochodzi do zerwania w miejscu przewężenia. Mówimy wtedy, że została przekroczona *granica wytrzymałości ciała*.

W tabeli 8.1 obok wartości modułów sprężystości podane są wartości granic sprężystości i wytrzymałości. Jeśli granice sprężystości i wytrzymałości mało się różnią, to mówimy, że ciało jest kruche. Jako przykład ciała kruchego często wymienia się szkło, pamiętać jednak należy, że w ścisłym, fizycznym znaczeniu szkło nie należy do ciał stałych: jest to ciało bezpostaciowe, a więc stanowi przechłodzoną ciecz.

Właściwości sprężyste ciał stałych zależą od temperatury, ale tych spraw dokładniej nie będziemy rozpatrywać.

Omawiane wyżej odkształcenia sprężyste powiązaliśmy z trzema modułami spręż-

żystości: modułem Younga E, modułem sprężystości objętościowej K oraz modułem sprężystości postaciowej G. Przy bliższych badaniach okazało się, że moduł E można powiązać z modułami K i G, wprowadzając dodatkowo współczynnik zwany *liczbą Poissona* μ. Sens fizyczny tego współczynnika wiąże się z wydłużeniem sprężystym. A mianowicie podczas rozciągania ciała pojawia się równoczesne zwężenie w kierunku prostopadłym do działającej siły. Stosunek względnego wydłużenia ε_t w kierunku działającej siły do względnego skrócenia $-\varepsilon_n$ w kierunku prostopadłym wyznacza liczbę Poissona

$$\mu = -\frac{\varepsilon_t}{\varepsilon_n}.$$

Podajemy bez dowodu związki między E, K, G i μ:

$$E = 2G(1+\mu),$$

$$E = 3K(1-2\mu).$$

Na zakończenie warto podkreślić, że przy stosowaniu niewielkich, lecz wielokrotnie (setki, a nawet tysiące razy) powtarzających się odkształceń, ciała stałe wykazują tzw. *zjawisko zmęczenia*. Polega ono na tym, że dany materiał może ulec uszkodzeniu (rozerwaniu, złamaniu) nawet wtedy, gdy stosowane ciśnienia są znacznie mniejsze od granicy wytrzymałości, jeśli tylko stosowane są one wielokrotnie. Im większe jest zastosowane ciśnienie, tym mniejsza liczba wywołanych odkształceń wystarcza do spowodowania uszkodzenia.

8.3. Energia potencjalna sprężysta

Siły zewnętrzne powodujące odkształcenia ciała działają zawsze na pewnej drodze, a więc zawsze wykonują pewną pracę. Praca ta zostaje zmagazynowana w ciele pod postacią energii potencjalnej sprężystej, która w odpowiednich warunkach może być znów przekształcona na pracę lub na inny rodzaj energii.

Dla przykładu obliczmy energię potencjalną całkowitą oraz gęstość energii potencjalnej (tzn. stosunek energii potencjalnej całkowitej E_p do objętości ciała V) w przypadku sprężystego pręta o długości l_0, który został wydłużony o $\Delta l = x_1$.

Energię potencjalną całkowitą obliczamy korzystając ze wzoru na pracę (3.3):

$$E_p = W = \int_0^{x_1} Fd(\Delta l) = \int_0^{x_1} F dx.$$

Symbol x w powyższym wyrażeniu przedstawia zmienne wydłużenie bezwzględne Δl. Granicami zmienności x są zero i x_1.

Uwzględniając wzór (8.4) otrzymujemy

$$E_p = \int_0^{x_1} ES\frac{x}{l_0}\,dx,$$

a po scałkowaniu

$$E_p = \frac{ES}{l_0}\frac{x_1^2}{2}.$$

Gęstość energii potencjalnej w wyraża się wzorem

$$w = \frac{E_p}{V} = \frac{E_p}{Sl_0} = \frac{Ex_1^2}{2l_0^2}.$$

Uwzględniając zależność

$$\frac{x_1}{l_0} = \frac{\Delta l}{l_0} = \varepsilon$$

otrzymujemy ostatecznie

$$w = \frac{E\varepsilon^2}{2} = \frac{\sigma^2}{2E}.$$

Analogiczne wzory otrzymalibyśmy badając odkształcenie objętościowe lub postaciowe, a mianowicie

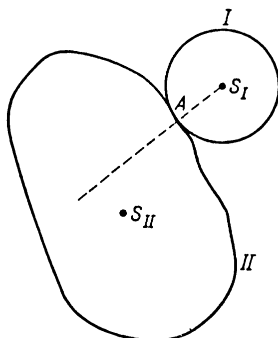
$$w = \frac{\sigma^2}{2K} \quad \text{oraz} \quad w = \frac{\tau^2}{2G}.$$

W każdym z wymienionych przypadków *gęstość energii potencjalnej jest proporcjonalna do kwadratu naprężenia* (normalnego σ lub stycznego τ) i *odwrotnie proporcjonalna do odpowiedniego modułu sprężystości.*

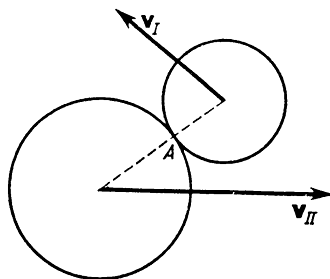
8.4. Ogólna charakterystyka zderzeń ciał

Omawiając właściwości ciał stałych odkształcalnych warto nieco uwagi poświęcić zjawisku zderzeń. Zaczniemy od krótkich uwag o charakterze klasyfikacyjnym.

Zderzenia dzielimy na *centralne* i *niecentralne. Zderzenie* jest *centralne*, jeżeli normalna do powierzchni wystawiona w punkcie A zetknięcia zderzających się ciał prze-



Rys. 8.5



Rys. 8.6

chodzi przez środek masy ciała. Rys. 8.5 przedstawia zderzenie centralne dla ciała I i jednocześnie niecentralne dla ciała II. Punkty S_{I} i S_{II} oznaczają odpowiednio środki masy obu ciał. Zderzenia kul jednorodnych są zawsze centralne.

Inny podział zderzeń rozróżnia *zderzenia proste i skośne. Zderzenie* jest *proste*, jeżeli kierunki normalnych wystawionych jak poprzednio w punkcie *A* zetknięcia ciał pokrywają się z kierunkami prędkości. Rys. 8.6 przedstawia zderzenie skośne dwóch kul. W przypadku kul zderzenia są proste wtedy, gdy środki kul poruszają się po tej samej prostej, np. gdy jedna kula dogania drugą lub obie przesuwają się wprost ku sobie.

Biorąc pod uwagę charakter odkształceń zachodzących w ciałach podczas zderzeń podzielimy je na *sprężyste* i *niesprężyste*. Mimo że w przyrodzie nie występują ani ciała doskonale sprężyste, ani ciała doskonale niesprężyste, wygodnie jest posługiwać się pojęciem *zderzeń doskonale sprężystych* i *doskonale niesprężystych*. Jest to znowu idealizacja zjawisk upraszczająca rozważania matematyczne. Trzeba jednak pamiętać, że wyniki otrzymane przy takich upraszczających założeniach mają zastosowanie tylko w sposób mniej lub więcej przybliżony do zjawisk rzeczywiście zachodzących w przyrodzie. Tak np. zderzenia dwóch ciał z kości słoniowej lub dobrej stali są prawie doskonale sprężyste. Zderzenia ciał z wosku, gliny lub ołowiu podobne są do zderzeń doskonale niesprężystych.

Spróbujmy krótko wyjaśnić zasadniczą różnicę między zderzeniami doskonale sprężystymi i doskonale niesprężystymi, dokonując analizy przebiegu tych procesów. W obu rodzajach zderzeń można wyróżnić tę samą fazę początkową procesu, podczas której względna prędkość zderzających się ciał spada do zera. Towarzyszy temu zmniejszanie się łącznej energii kinetycznej obu ciał. Ten ubytek energii kinetycznej zamienia się na pracę i tu zaznacza się różnica w przebiegu obu procesów. W przypadku zderzenia doskonale sprężystego mamy do czynienia z pracą odkształcenia sprężystego, a więc z pracą pokonywania sił zachowawczych. W przypadku zderzenia doskonale niesprężystego jest to praca trwałego odkształcenia. Praca pokonywania sił zachowawczych, działających przy zderzeniu doskonale sprężystym, może nam być w całości zwrócona: ciała wyzwalają się stopniowo z odkształcenia i po krótkim czasie odrywają się od siebie odbiegając z pewnymi prędkościami. W tym procesie część początkowej całkowitej energii kinetycznej tylko przejściowo zamienia się na energię potencjalną sprężystą. Wyzwalaniu się z odkształcenia towarzyszy zmiana energii potencjalnej na kinetyczną, tak że w końcowym etapie zderzenia znowu mamy do czynienia tylko z energią kinetyczną. Wobec tego, że w rozważanym odosobnionym układzie działają wyłącznie siły zachowawcze, nie ma strat energii, a zatem suma energii kinetycznych ciał przed zderzeniem równa się sumie energii kinetycznych po zderzeniu.

Praca trwałego odkształcenia wykonana w czasie zderzenia doskonale niesprężystego nie może być nam zwrócona. Po zmniejszeniu się do zera prędkości względnej zderzających się ciał (tzn. gdy oba ciała mają wspólną prędkość) przestają one na siebie oddziaływać i poruszają się dalej jako jedna bryła. Tym razem energia kinetyczna powstałej bryły jest mniejsza od sumy energii kinetycznych obu ciał przed zderzeniem.

W punktach 8.4.1 i 8.4.2 podamy analizę matematyczną zderzeń centralnych doskonale sprężystych i doskonale niesprężystych.

8.4.1. *Zderzenia kul doskonale sprężystych.* Rozpatrzmy dokładniej zderzenia proste dwóch jednorodnych, doskonale sprężystych kul o masach m_1 i m_2 oraz o prędkościach przed zderzeniem v_1 i v_2. Chcemy obliczyć prędkości u_1 i u_2 obu kul po zderzeniu.

Powiedzieliśmy już, że zderzenie doskonale sprężyste charakteryzuje to, że energia kinetyczna przed zderzeniem równa jest energii kinetycznej po zderzeniu:

$$\tfrac{1}{2}m_1v_1^2+\tfrac{1}{2}m_2v_2^2 = \tfrac{1}{2}m_1u_1^2+\tfrac{1}{2}m_2u_2^2. \tag{8.7}$$

Zderzające się kule można traktować jako układ odosobniony, w którym działają tylko siły wewnętrzne. Obowiązuje więc zasada zachowania pędu:

$$m_1v_1+m_2v_2 = m_1u_1+m_2u_2. \tag{8.8}$$

Z wzoru (8.7) otrzymujemy

$$m_1(v_1^2-u_1^2) = m_2(u_2^2-v_2^2), \tag{8.9}$$

a z wzoru (8.8) wynika, że

$$m_1(v_1-u_1) = m_2(u_2-v_2). \tag{8.10}$$

Dzieląc stronami ostatnie równania otrzymujemy

$$v_1+u_1 = u_2+v_2, \tag{8.11}$$

skąd

$$u_2 = v_1+u_1-v_2. \tag{8.12}$$

Podstawiamy do wzoru (8.10):

$$m_1(v_1-u_1) = m_2(v_1+u_1-2v_2),$$

$$u_1 = \frac{2m_2v_2+(m_1-m_2)v_1}{m_1+m_2}. \tag{8.13}$$

Wracając do równania (8.12) znajdujemy

$$u_2 = \frac{(v_1-v_2)(m_1+m_2)+2m_2v_2+(m_1-m_2)v_1}{m_1+m_2},$$

$$u_2 = \frac{2m_1v_1+(m_2-m_1)v_2}{m_1+m_2}. \tag{8.14}$$

Przejdźmy do szczególnych przypadków zderzeń prostych kul doskonale sprężystych.

1. Niech $m_1 = m_2$. Wtedy ze wzorów (8.13) i (8.14) wynika

$$u_1 = v_2 \quad \text{i} \quad u_2 = v_1.$$

Przy zderzeniu prostym dwie kule o równych masach wymieniają wzajemnie prędkości.

2. Jeśli druga kula przed zderzeniem jest nieruchoma, czyli $v_2 = 0$, to

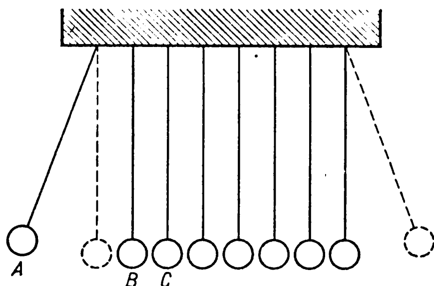
$$u_1 = \frac{m_1-m_2}{m_1+m_2}v_1, \quad \text{a} \quad u_2 = \frac{2m_1v_1}{m_1+m_2}.$$

Jeśli równocześnie $m_1 = m_2$, wówczas

$$u_1 = 0, \quad u_2 = v_1.$$

Przypadek ten ilustruje doświadczenie przedstawione na rys. 8.7. Widzimy tam kilka kul stalowych o równych masach, zawieszonych na nitkach o tej samej długości. Kule stykają się ze sobą. Kulę *A* wyprowadzamy z położenia równowagi, zachowując jej środek w płaszczyźnie pionowej przechodzącej przez środki kul pozostałych. Puszczona swobodnie kula opada, zderza się z kulą *B*, przekazuje jej swą energię, a sama się za-

trzymuje. Kula B przekazuje uzyskaną energię kuli C itd. Odskakuje tylko kula ostatnia z tą samą prędkością, z jaką kula A uderzyła o kulę B. Kula ostatnia wznosi się na tę samą wysokość, na jaką początkowo została podniesiona kula A (przy założeniu doskonałej sprężystości stali i braku wszelkich oporów).



Rys. 8.7

3. Gdy druga kula ma masę znacznie większą od pierwszej i jest nieruchoma, czyli gdy $m_1 \ll m_2$ i $v_2 = 0$, wtedy

$$u_1 = \frac{\frac{m_1}{m_2} - 1}{\frac{m_1}{m_2} + 1} v_1, \quad \text{a} \quad u_2 = \frac{2\frac{m_1}{m_2}}{\frac{m_1}{m_2} + 1} v_1 .$$

Jeśli założymy, że masa $m_2 \to \infty$, to

$$\lim_{m_2 \to \infty} \frac{m_1}{m_2} = 0, \quad \text{a więc} \quad u_1 \to -v_1, \quad u_2 \to 0.$$

Po zderzeniu kula o bardzo dużej masie pozostaje nadal nieruchoma, a kula o masie mniejszej porusza się z prędkością tą samą co do wartości, lecz przeciwnie skierowaną.

Ten sam wynik odnosi się do odbicia prostopadłego od masywnej nieruchomej podstawy lub ściany. Stąd wniosek, że kulka stalowa puszczona swobodnie w próżni odskakuje od podstawy stalowej prawie na tę samą wysokość, z jakiej spada.

4. Jeśli $m_1 \gg m_2$, a równocześnie $v_2 = 0$, to

$$u_1 = \frac{1 - \frac{m_2}{m_1}}{1 + \frac{m_2}{m_1}} v_1, \quad \text{a} \quad u_2 = \frac{2v_1}{1 + \frac{m_2}{m_1}} .$$

Jeśli założymy, że $m_1 \to \infty$, to

$$\lim_{m_1 \to \infty} \frac{m_2}{m_1} = 0,$$

a wtedy

$$u_1 \to v_1, \quad u_2 \to 2v_1 .$$

Innymi słowy, po zderzeniu kuli o bardzo dużej masie z nieruchomą kulką o masie

bardzo małej, praktycznie biorąc kula duża zachowuje swą prędkość pierwotną, a kulka mała odskakuje z prędkością dwa razy większą od prędkości kuli dużej.

Interesujące mogą być dane dotyczące czasu zderzeń, tzn. czasu, w ciągu którego ciała zderzające się działają na siebie dużymi siłami. Na przykład czas trwania zderzenia dwóch kulek stalowych o średnicy 0,05 m, poruszających się z prędkością względną 0,02 m/s, wynosi 0,00038 s. Obliczono, że gdyby obie kule miały rozmiary kuli ziemskiej, czas zderzenia wynosiłby około 27 godzin.

Zderzeniom sprężystym poświęciliśmy dużo uwagi nie tylko dlatego, że są piękną ilustracją stosowalności poznanych poprzednio zasad zachowania energii mechanicznej i pędu, lecz także dlatego, że często będziemy się do nich odwoływać, gdy będzie mowa o zderzeniach atomowych i cząsteczkowych (por. np. rozdz. 16). Od razu jednak podkreślimy, że nie wszystkie zderzenia atomowe mają charakter sprężysty: często zdarzają się takie, przy których uderzony atom lub cząsteczka absorbują część energii kinetycznej ciała uderzającego zużywając ją na wzbudzenie lub jonizację (por. np. rozdz. 20 lub 30).

8.4.2. *Zderzenia doskonale niesprężyste.* Rozważymy ten rodzaj zderzeń na przykładzie dwóch doskonale niesprężystych ciał o masach m_1 i m_2 oraz o prędkościach przed zderzeniem v_1 i v_2. Niech obie prędkości mają te same kierunki i v_1 niech będzie większe od v_2, czyli niech ciało pierwsze dogania drugie. Po zderzeniu, jak już wiemy, następuje trwałe odkształcenie obu ciał i biegną one dalej jako jedna bryła z prędkością u.

W czasie tego zderzenia nie działają w układzie odosobnionym siły zachowawcze, a zatem nie stosuje się zasada zachowania energii mechanicznej. Stosuje się natomiast zasada zachowania pędu

$$m_1 v_1 + m_2 v_2 = (m_1 + m_2)u.$$

Stąd prędkość wspólna obu ciał po zderzeniu równa się

$$u = \frac{m_1 v_1 + m_2 v_2}{m_1 + m_2}. \tag{8.15}$$

Znając energię kinetyczną obu ciał przed zderzeniem, jak również energię kinetyczną bryły utworzonej w wyniku zderzenia, można obliczyć stratę energii kinetycznej $\Delta E_k = E$, przekształconą na inne postacie energii:

$$E = \frac{m_1 v_1^2}{2} + \frac{m_2 v_2^2}{2} - \frac{(m_1 + m_2)u^2}{2}.$$

Po uwzględnieniu (8.15) otrzymuje się ostatecznie

$$E = \frac{1}{2} \frac{m_1 m_2}{m_1 + m_2} (v_1 - v_2)^2.$$

Iloczyn obu zderzających się mas podzielony przez ich sumę przedstawia tzw. *masę zredukowaną układu.* Różnica $(v_1 - v_2)$ jest prędkością względną. A zatem *ubytek energii kinetycznej przekształcony w czasie prostego doskonale niesprężystego zderzenia na inne rodzaje energii jest proporcjonalny do masy zredukowanej układu oraz kwadratu prędkości względnej.*

Właściwości cieczy i gazów

W rozdziale niniejszym zajmować się będziemy płynami, tzn. cieczami i gazami. Zestawimy podstawowe wiadomości z hydrostatyki i aerostatyki, jak również hydrodynamiki i aerodynamiki. Rozpoczniemy od rozważań hydrostatycznych.

9.1. Podstawowe wiadomości z hydrostatyki

9.1.1. *Ogólna charakterystyka cieczy.* Ciecz nie ma określonego kształtu (postaci), lecz przyjmuje kształt naczynia, do którego jest wlana. W związku z tym moduł sprężystości postaciowej G cieczy równa się zeru.

Ciecz ma powierzchnię swobodną. Gdy ciecz jest poddana wyłącznie działaniu siły ciężkości, powierzchnia swobodna cieczy jest pozioma. Odchylenia od poziomu występują np. przy ściankach naczynia lub w pobliżu powierzchni jakiegoś ciała zanurzonego w cieczy. Jest to wynik działania sił międzycząsteczkowych, zwanych siłami przylegania. Wpływ tych sił zaznacza się najwyraźniej w naczyniach o niewielkich przekrojach, a więc przede wszystkim w rurkach włoskowatych. Siła działająca na cząsteczkę cieczy, nieruchomą w stosunku do cząsteczek pozostałych, w dowolnym punkcie powierzchni swobodnej jest skierowana prostopadle do powierzchni. Gdyby bowiem siła była skierowana ukośnie względem powierzchni, to posiadałaby składową normalną i styczną, a składowa styczna wywoływałaby ruch danej cząsteczki względem cząsteczek sąsiednich.

P y t a n i e. Wyjaśnić dlaczego powierzchnia swobodna cieczy wlanej do walcowatego naczynia obracającego się dokoła osi pionowej (rys. 9.1) ma kształt paraboloidy obrotowej? Jakie siły działają na każdą cząsteczkę cieczy na jej powierzchni swobodnej? Jak skierowana jest siła wypadkowa względem powierzchni swobodnej?

Ciecze odznaczają się bardzo małą ściśliwością, tzn. trzeba bardzo dużych ciśnień, aby w sposób nieznaczny zmniejszyć objętość cieczy. Podczas badań sprężystości objętościowej cieczy stosowano ciśnienia dochodzące do ok. $15 \cdot 10^9$ N/m^2.

Prawo Hooke'a w odniesieniu do cieczy piszemy często w postaci

$$\frac{1}{K}p = -\frac{\Delta V}{V_0}.$$

Współczynnik $1/K$ nazywamy współczynnikiem ściśliwości cieczy. Wartość licz-

Właściwości cieczy i gazów

W rozdziale niniejszym zajmować się będziemy płynami, tzn. cieczami i gazami. Zestawimy podstawowe wiadomości z hydrostatyki i aerostatyki, jak również hydrodynamiki i aerodynamiki. Rozpoczniemy od rozważań hydrostatycznych.

9.1. Podstawowe wiadomości z hydrostatyki

9.1.1. *Ogólna charakterystyka cieczy.* Ciecz nie ma określonego kształtu (postaci), lecz przyjmuje kształt naczynia, do którego jest wlana. W związku z tym moduł sprężystości postaciowej G cieczy równa się zeru.

Ciecz ma powierzchnię swobodną. Gdy ciecz jest poddana wyłącznie działaniu siły ciężkości, powierzchnia swobodna cieczy jest pozioma. Odchylenia od poziomu występują np. przy ściankach naczynia lub w pobliżu powierzchni jakiegoś ciała zanurzonego w cieczy. Jest to wynik działania sił międzycząsteczkowych, zwanych siłami przylegania. Wpływ tych sił zaznacza się najwyraźniej w naczyniach o niewielkich przekrojach, a więc przede wszystkim w rurkach włoskowatych. Siła działająca na cząsteczkę cieczy, nieruchomą w stosunku do cząsteczek pozostałych, w dowolnym punkcie powierzchni swobodnej jest skierowana prostopadle do powierzchni. Gdyby bowiem siła była skierowana ukośnie względem powierzchni, to posiadałaby składową normalną i styczną, a składowa styczna wywoływałaby ruch danej cząsteczki względem cząsteczek sąsiednich.

P y t a n i e. Wyjaśnić dlaczego powierzchnia swobodna cieczy wlanej do walcowatego naczynia obracającego się dokoła osi pionowej (rys. 9.1) ma kształt paraboloidy obrotowej? Jakie siły działają na każdą cząsteczkę cieczy na jej powierzchni swobodnej? Jak skierowana jest siła wypadkowa względem powierzchni swobodnej?

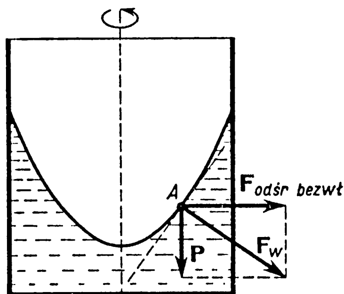
Ciecze odznaczają się bardzo małą ściśliwością, tzn. trzeba bardzo dużych ciśnień, aby w sposób nieznaczny zmniejszyć objętość cieczy. Podczas badań sprężystości objętościowej cieczy stosowano ciśnienia dochodzące do ok. $15 \cdot 10^9$ N/m^2.

Prawo Hooke'a w odniesieniu do cieczy piszemy często w postaci

$$\frac{1}{K}p = -\frac{\Delta V}{V_0}.$$

Współczynnik $1/K$ nazywamy współczynnikiem ściśliwości cieczy. Wartość licz-

bowa $1/K$ odpowiada $\Delta V/V_0$, przy założeniu, że p równa się 1. A zatem *współczynnik ściśliwości cieczy* jest to ułamek wskazujący, o jaką część objętości początkowej zmienia się objętość danej cieczy podczas zmiany ciśnienia o jednostkę. Oczywiście wartość liczbowa $1/K$ zależy od tego, w jakich jednostkach mierzone jest ciśnienie. W układzie SI współczynnik ściśliwości wyraża się w $(N/m^2)^{-1}$.



Rys. 9.1

Wartość współczynnika ściśliwości zależy od ciśnienia, jakiemu ciecz jest poddana. Ze wzrostem ciśnienia wartość współczynnika maleje. Jaki sens fizyczny zawarty jest w tym stwierdzeniu?

Wartość współczynnika ściśliwości danej cieczy zależy też od temperatury. W przypadku wody zależność jest taka, że podczas ogrzewania od zera do 100°C współczynnik ściśliwości początkowo maleje, dochodzi do pewnego minimum, a następnie rośnie.

Tabela 9.1

Współczynniki ściśliwości cieczy

Ciecz	Temperatura °C	Ciśnienie 10^5 N/m²	Współczynnik ściśliwości 10^{-11} m²/N
alkohol etylowy	0	1–50	96
	0	100–200	85
	0	300–400	73
	28	150–200	86
	100	150–200	168
eter	12	0,4–17	163
gliceryna	14,9	1–10	22
rtęć	15	100–200	3,76
woda	0	1–25	52,5
	0	25–50	51,6
	0	100–200	49,2
	10	1–25	50
	10	25–50	49,2
	10	100–200	46,1
	20	1–25	49,1
	20	25–50	47,6
	20	100–200	44,2

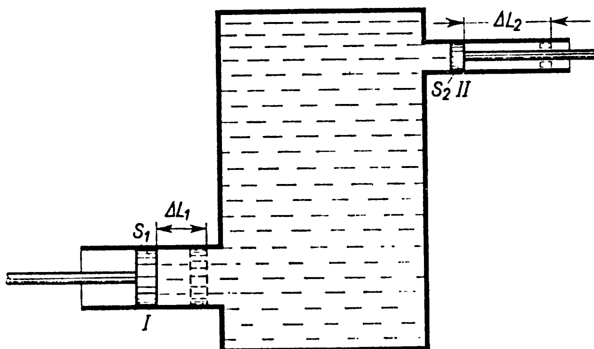
W tabeli 9.1 podane są wartości współczynników ściśliwości kilku cieczy w różnych warunkach. Z tabeli widać, że współczynniki te wyrażają się bardzo małymi ułamkami. To uzasadnia wprowadzenie pojęcia *cieczy doskonałej* jako cieczy *całkowicie nieściśliwej*. Ciecz doskonała pozostająca w spoczynku spełnia zatem warunek $1/K = 0$. *Ciecz doskonała będąca w ruchu spełnia jeszcze* jeden warunek — *warunek braku lepkości*, o czym będziemy mówili dalej.

Ciecz doskonała w przyrodzie nie istnieje. Mamy więc znowu do czynienia z ujęciem modelowym, z idealizacją zjawisk upraszczającą rozważania matematyczne, lecz prowadzącą do wyników mniej lub bardziej przybliżonych. Na razie zajmować się będziemy właściwościami cieczy doskonałej, pozostającej w równowadze statycznej.

9.1.2. *Prawo Pascala.* Prawo Pascala odnosi się do cieczy nieściśliwej i ma zastosowanie wtedy, gdy ciśnienie wynikające z własnego ciężaru cieczy równa się zeru (np. w stanie nieważkości w sztucznym satelicie) lub jest niewielkie w porównaniu z ciśnieniem zewnętrznym, tak że można je zaniedbać. Jeśli przyjmiemy dla takiej cieczy nazwę *cieczy nieważkiej*, to treść prawa Pascala można będzie ująć następująco: *w cieczy nieściśliwej i nieważkiej ciśnienie zewnętrzne rozchodzi się we wszystkich kierunkach jednakowo.*

Wyobraźmy sobie naczynie wypełnione cieczą nieściśliwą i nieważką, mające dwa otwory boczne zamknięte tłokami o polach powierzchni S_1 i S_2 (rys. 9.2). Na tłok I działa ciśnienie zewnętrzne p_1 przesuwające ten tłok o odległość ΔL_1. Zostaje przy tym wykonana praca W_1 równa

$$W_1 = p_1 S_1 \Delta L_1 .$$



Rys. 9.2

Praca ta jest bez strat przekazana przez ciecz tłokowi II, który ulega przesunięciu o ΔL_2 pod działaniem ciśnienia p_2 ze strony cieczy. A zatem

$$W_2 = p_2 S_2 \Delta L_2 .$$

Z równości obu prac wynika zależność

$$p_1 S_1 \Delta L_1 = p_2 S_2 \Delta L_2 .$$

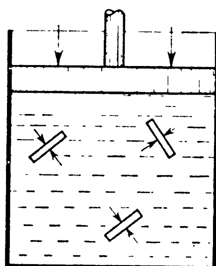
Ale iloczyny $S_1 \Delta L_1$ i $S_2 \Delta L_2$ wyrażają objętości cieczy zawarte między początkowymi i końcowymi położeniami obu tłoków. Wobec nieściśliwości cieczy te objętości są sobie równe, a zatem

$$p_1 = p_2 .$$

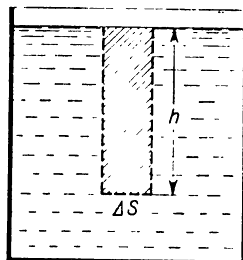
Ciśnienie zewnętrzne p_1 bez zmiany wartości zostało przekazane na tłok II; ponieważ jednak jego położenie było dowolne, więc można wyciągnąć wniosek, że ciśnienie zewnętrzne p_1 rozeszło się w cieczy w ten

sposób, że na każdym elemencie powierzchni pomyślanym w cieczy ma ono stałą wartość (rys. 9.3). Prostopadłość kierunku ciśnienia względem elementu powierzchni wynika z założenia o równowadze cieczy.

Z prawem Pascala wiąże się działanie prasy hydraulicznej. Jest to urządzenie służące do otrzymywania dużych sił (dużego parcia). Zasada działania prasy hydraulicznej znana jest z kursu szkoły średniej.



Rys. 9.3



Rys. 9.4

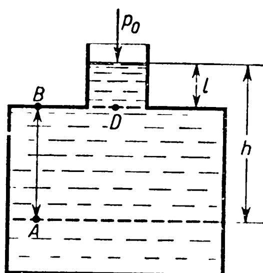
9.1.3. *Ciśnienie hydrostatyczne.* Odrzućmy obecnie warunek nieważkości wprowadzony przy prawie Pascala i przejdźmy do ciśnienia wywieranego przez ciecz dzięki jej ciężarowi, czyli do tzw. *ciśnienia hydrostatycznego*. Wydzielmy w cieczy o gęstości ϱ na głębokości h pod powierzchnią swobodną element powierzchni ΔS (rys. 9.4). Ciśnienie hydrostatyczne p wywierane na ten element powierzchni można obliczyć z następującej zależności:

$$p = \frac{mg}{\Delta S} = \frac{\Delta S h \varrho g}{\Delta S}.$$

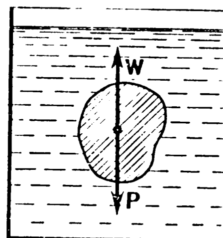
Ciśnienie hydrostatyczne wyraża się więc wzorem

$$p = h\varrho g. \tag{9.1}$$

Jest ono proporcjonalne do gęstości cieczy i głębokości zanurzenia liczonej względem powierzchni swobodnej.



Rys. 9.5



Rys. 9.6

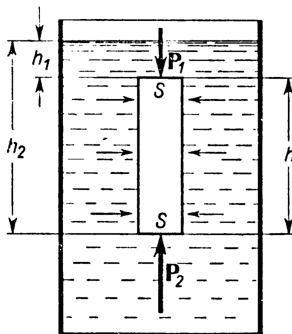
Zgodnie z tym, co powiedzieliśmy, na element powierzchni zawierający punkt A (rys. 9.5) działa ciśnienie hydrostatyczne $(AB+l)\varrho g$, a na element zawierający punkt B — ciśnienie $l\varrho g$, takie samo jak na element zawierający punkt D. Na wszystkich ele-

mentach powierzchni, znajdujących się na tym samym poziomie w cieczy pozostającej w spoczynku, istnieje to samo ciśnienie. Kojarzy się z tym znane prawo równowagi cieczy w naczyniach połączonych: ciecz w naczyniach połączonych pozostaje w spoczynku, czyli jest w równowadze, gdy ciśnienia na wspólnym poziomie tej samej cieczy w różnych naczyniach są jednakowe.

Jeżeli ciśnienie zewnętrzne, np. atmosferyczne, działające na powierzchnię swobodną cieczy w naczyniu wynosi p_0, to ciśnienie całkowite na głębokości h mierzonej od powierzchni swobodnej wynosi

$$p = p_0 + h\varrho g. \tag{9.2}$$

Fakt, że ciśnienie hydrostatyczne na określonym poziomie w danej cieczy zależy jedynie od jej gęstości, decyduje o występowaniu tzw. paradoksu hydrostatycznego. Przez ***paradoks hydrostatyczny*** rozumiemy niezależność parcia i ciśnienia, wywieranych przez ciecz na dno naczyń, od kształtu tych naczyń, jeśli tylko powierzchnie ich den i wysokości słupów wlanej cieczy są jednakowe.



Rys. 9.7

9.1.4. *Prawo Archimedesa.* Wyobraźmy sobie, że wyodrębniliśmy w cieczy znajdującej się w spoczynku pewną objętość (obszar zakreślony na rys. 9.6). Na ciecz w tej objętości działa pionowo w dół siła ciężkości **P** przyłożona w środku ciężkości rozważanej masy cieczy. Jeśli całość cieczy pozostaje w spoczynku, to znaczy, że wspomniana siła ciężkości jest zrównoważona przez oddziaływanie wypadkowe otaczającej cieczy równe też **P**, lecz skierowane pionowo w górę. Niech ciecz w rozważanej objętości skrzepnie zachowując niezmienioną gęstość. W warunkach równowagi nic nie ulegnie zmianie. Zastąpmy teraz w myśli wyodrębnioną objętość cieczy przez bryłę z dowolnego materiału o tym samym kształcie i rozmiarach. Oddziaływanie ze strony cieczy pozostanie niezmienione. Ciało zanurzone będzie podlegało wypadkowemu parciu skierowanemu do góry, zwanemu *wyporem*. Wypór W jest równy ciężarowi takiej ilości cieczy, która ma objętość równą objętości ciała zanurzonego. Innymi słowy, *wypór, jakiemu podlega ciało zanurzone w cieczy, równa się ciężarowi cieczy wypartej przez to ciało.* Jest to *treść prawa Archimedesa* odkrytego już 250 lat przed naszą erą.

Do prawa Archimedesa doprowadza następujące rozumowanie. Wyodrębniamy w cieczy walec o wysokości h i o polu podstawy S. Górna podstawa znajduje się na głębokości h_1 (rys. 9.7), dolna — na głę-

bokości h_2 pod powierzchnią swobodną cieczy. Na obie podstawy działają ciśnienia hydrostatyczne odpowiednio równe $p_1 = h_1 \varrho g$ i $p_2 = h_2 \varrho g$, prostopadłe do obu podstaw, skierowane: p_1 — pionowo w dół, p_2 — pionowo w górę. Odpowiadają im wartości parcia $P_1 = p_1 S$ i $P_2 = p_2 S$. Z ich równoczesnego występowania wynika parcie wypadkowe F równe

$$F = (h_2 - h_1)\varrho g S.$$

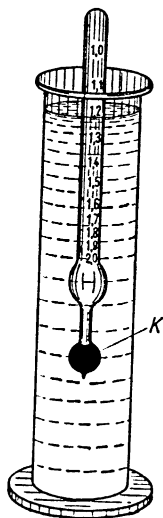
Ciśnienia hydrostatyczne działające na ściany boczne rosną wraz z głębokością, ale na każdej głębokości wzajemnie się równoważą. Znalezione parcie F jest zatem parciem wypadkowym, wynikającym z działania wszystkich ciśnień hydrostatycznych na ściany walca, a więc jest wyporem W. Wypór pozostanie niezmieniony, gdy walec wyodrębniony z cieczy zastąpimy walcem o tych samych rozmiarach z innego materiału

$$W = (h_2 - h_1)\varrho g S.$$

Ponieważ $(h_2 - h_1) = h$ jest wysokością walca, więc łatwo sprawdzić, że iloczyn po prawej stronie równania przedstawia ciężar P_{cw} cieczy wypartej przez walec. A zatem

$$W = P_{cw}. \tag{9.3}$$

Podobne rozumowanie można przeprowadzić dla bryły dowolnego kształtu, rozkładając ciśnienia hydrostatyczne na składowe poziome i pionowe. Wynik ostateczny jest ten sam.



Rys. 9.8

Z prawa Archimedesa wynikają *warunki pływania ciał*. Jeśli wypór W jest większy od ciężaru P_{cz} ciała zanurzonego — ciało pływa, jeśli jest mniejszy — ciało tonie, jeśli zaś $W = P_{cz}$ — ciało utrzymuje się w równowadze na dowolnej głębokości wewnątrz cieczy (przy założeniu nieściśliwości zarówno cieczy, jak i ciała zanurzonego). Ponieważ W i P_{cz} oznaczają ciężary cieczy i ciała zanurzonego o tej samej objętości, więc warunki pływania ciał jednorodnych można uzależnić od gęstości cieczy i ciała zanurzonego. A mianowicie:

$\varrho_{\text{ciała zanurzonego}} < \varrho_{\text{cieczy}}$ — ciało pływa,

$\varrho_{\text{ciała zanurzonego}} > \varrho_{\text{cieczy}}$ — ciało tonie,

$\varrho_{\text{ciała zanurzonego}} = \varrho_{\text{cieczy}}$ — ciało utrzymuje się na dowolnej głębokości w cieczy.

Oczywiście, analogiczne zależności obowiązują w odniesieniu do ciężarów właściwych ciała zanurzonego i cieczy.

Prawo Archimedesa służy często do obliczania ciężarów właściwych ciał stałych i ciekłych. Zostało to uwzględnione w pytaniu 7 i 8 na końcu rozdziału.

Do szybkiego wyznaczania ciężarów właściwych cieczy służy przyrząd zwany *areometrem* (rys. 9.8). Dolna część areometru — bańka K — wypełniona jest ciałem o dużym ciężarze właściwym, np. śrutem ołowianym. Górna część — w postaci wąskiej rurki — zaopatrzona jest w podziałkę podającą ciężary właściwe cieczy. Pomiar opiera się na prawie Archimedesa: areometr zanurza się w każdej cieczy tak głęboko, by jego własny ciężar P_a był zrównoważony przez wypór ze strony cieczy. A zatem przy każdym zanurzeniu spełniona jest zależność

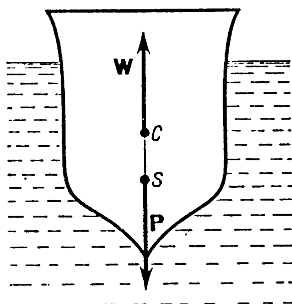
$$P_a = W = \gamma_c V_{zan} = \text{const},$$

gdzie γ_c oznacza ciężar właściwy cieczy, V_{zan} — objętość zanurzonej części areometru.

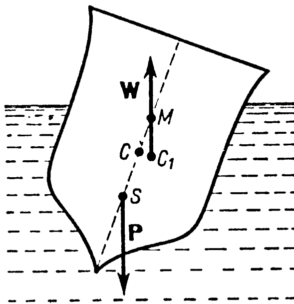
Liczby zaznaczone na skali areometru rosną w miarę przesuwania się ku dołowi, gdyż większym ciężarom właściwym cieczy odpowiadają mniejsze objętości zanurzenia.

P y t a n i e. Dlaczego odstępy między podziałkami areometru dotyczącymi większych ciężarów właściwych maleją?

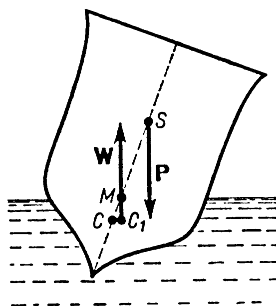
Wróćmy jeszcze do warunków pływania ciał na przykładzie statków, jachtów itp. (rys. 9.9). Siła ciężkości **P** statku przyłożona jest w środku ciężkości S, siła wyporu **W** — w punkcie C, zwanym *środkiem ciśnienia*, stanowiącym środek ciężkości tej masy wody, która by wypełniła objętość odpowiadającą zanurzonej części statku. Gdy (jak na rys. 9.9) punkty S i C leżą na jednym pionie, to zarówno siła wypadkowa, jak i moment wypadkowy równają się zeru i statek zachowuje równowagę. Przy przechyleniu się statku na bok (rys. 9.10) środek ciężkości S nie zmienia swego położenia względem statku, natomiast środek ciśnienia ulega przesunięciu z punktu C leżącego na osi statku do punktu C_1. Wypór **W** i ciężar statku **P** stanowią wtedy parę sił mającą moment obrotowy. W przypadku przedstawionym na rys. 9.10 ten moment obrotowy wywołuje powrót statku przez obrót do położenia poprzedniego. Para sił, taka jak na rys. 9.11, spowodowałaby przewrócenie się statku.



Rys. 9.9



Rys. 9.10



Rys. 9.11

Stałość równowagi statku można charakteryzować za pomocą położenia tzw. metacentrum. *Metacentrum* — to taki punkt M, w którym pion przeprowadzony przez C_1 (przy statku przechylonym) przecina kierunek pionu pierwotnego przechodzącego przez C i S. Jeśli metacentrum przypada powyżej środka ciężkości statku, to statek po przechyleniu wraca do swego położenia normalnego, które jest położeniem równowagi trwałej. Jeśli punkt M wypada poniżej S, para sił wywołuje przewrócenie się statku.

Zrozumiałe jest teraz stosowanie kilów ołowianych, np. w jachtach wyścigowych, celem zapewnienia trwałości równowagi. Silne obciążenie dolnej części jachtu obniża tak dalece środek ciężkości, że nawet po położeniu się na burtę jacht wraca do położenia normalnego.

9.2. Ogólna charakterystyka gazów

Gazy, podobnie jak ciecze, nie mają sprężystości postaciowej, czyli moduł sprężystości postaciowej gazów $G = 0$. Gazy nie mają swobodnej powierzchni, lecz wypełniają całkowicie zajmowane naczynie. Mają one cechę rozprężliwości, tzn. dążą do

zajęcia jak największej objętości. Ciśnienie w gazie spoczywającym rozchodzi się we wszystkich kierunkach jednakowo, a więc do gazów stosuje się *prawo Pascala*. Gęstość gazów w warunkach normalnych jest nieznaczna w porównaniu z gęstością cieczy lub ciał stałych (por. tab. 6.1).

Gazy wykazują sprężystość objętościową: ich objętość można zmieniać w znacznych granicach. Każdej zmianie objętości towarzyszy zmiana ciśnienia panującego w gazie (zmiana *prężności gazu*). Zależność objętości gazu od jego ciśnienia ujmuje prawo Boyle'a–Mariotte'a. Prawo to odnosi się do przemian tzw. *izotermicznych*, tzn. do przemian zachodzących w stałej temperaturze.

Prawo Boyle'a–Mariotte'a mówi, że *iloczyn objętości i ciśnienia dla danej masy gazu jest wielkością stałą w stałej temperaturze*:

$$pV = \text{const}, \quad \text{gdy} \quad t = \text{const}. \tag{9.4}$$

Niech wzrost ciśnienia gazu o Δp powoduje zmniejszenie objętości o ΔV. Spełnione jest wtedy równanie

$$pV = (p+\Delta p)(V-\Delta V),$$

skąd

$$pV = pV+\Delta pV-p\Delta V-\Delta p\Delta V.$$

Jeśli zmiana ciśnienia Δp jest bardzo mała, to i ΔV jest bardzo małe. Wtedy nie popełnimy większego błędu, jeżeli zaniedbamy wyraz $\Delta p\Delta V$ jako wielkość małą drugiego rzędu. Ostatni wzór przyjmie wówczas postać przybliżoną

$$\Delta pV = p\Delta V, \quad \text{skąd} \quad \Delta p = p\frac{\Delta V}{V}.$$

Możemy powiedzieć, że zastosowane ciśnienie dodatkowe Δp jest proporcjonalne do wywołanego względnego przyrostu objętości $\Delta V/V$. Ta zależność odpowiada *prawu Hooke'a*. Modułem sprężystości objętościowej jest tym razem p:

$$K = p. \tag{9.5}$$

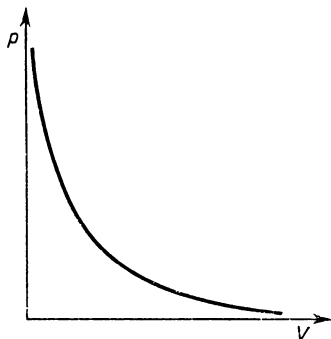
Stąd wniosek, że *moduł sprężystości objętościowej podczas przemian izotermicznych gazu równa się jego prężności*. Gaz ściśnięty w mniejszej objętości ma większą prężność, a tym samym ma większy moduł sprężystości objętościowej.

Badając gazy pod różnymi ciśnieniami w stałej temperaturze stwierdzamy, że prawo Boyle'a–Mariotte'a nie stosuje się dokładnie. Mimo to wprowadzono pojęcie gazu doskonałego, tzn. takiego, który dokładnie spełnia prawo Boyle'a–Mariotte'a. (Inne warunki, jakie spełniać musi gaz doskonały, omówione są w § 14.5) Dla takiego gazu zależność p od V w stałej temperaturze przedstawia się na wykresie jako hiperbola (rys. 9.12). Nadajemy jej w tym przypadku miano izotermy.

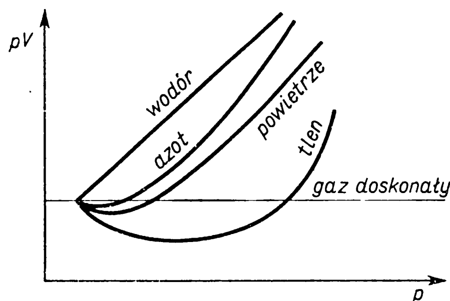
Odłóżmy na osi rzędnych wartości iloczynu pV, a na osi odciętych wartości p. Dla gazu doskonałego otrzymamy wtedy na wykresie (rys. 9.13) linię prostą równoległą do osi odciętych. Na tym samym wykresie zaznaczone są krzywe przedstawiające zależność pV od p dla kilku gazów rzeczywistych. Wyraźnie zaznaczają się odchylenia właściwości gazów rzeczywistych od właściwości gazu doskonałego. Odchylenia te są

mniejsze, gdy stosujemy niewielkie ciśnienia (mniejsze od 10^5 N/m²). Wykres 9.14 przedstawia przebieg zależności iloczynu pV od p dla jednego gazu (dwutlenku węgla) w różnych temperaturach zaznaczonych w skali Celsjusza przy poszczególnych krzywych. Odchylenia od właściwości gazu doskonałego są mniejsze w temperaturach wyższych.

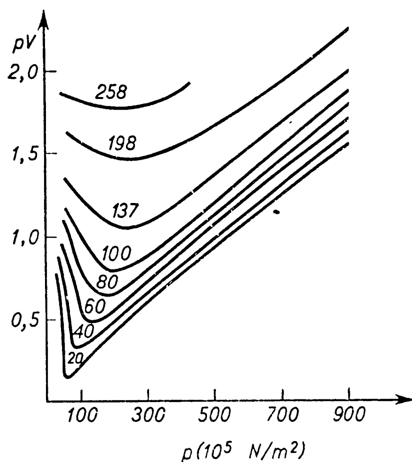
Na ogół do gazów rzeczywistych przy ciśnieniach do 10^6–10^7 N/m² można stosować prawo Boyle'a–Mariotte'a, pamiętać jednak należy, że opisuje ono zachowanie się gazu rzeczywistego tylko z pewnym przybliżeniem.



Rys. 9.12



Rys. 9.13



Rys. 9.14

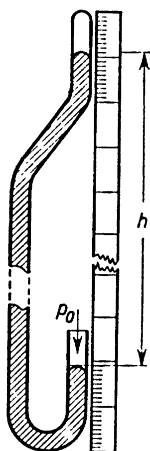
Do gazów stosuje się także *prawo Archimedesa. Ciało zanurzone w gazie doznaje wyporu skierowanego pionowo do góry, równego ciężarowi wypartego gazu.*

Ziemia otoczona jest grubą warstwą powietrza, zwaną atmosferą ziemską. Ponieważ ciała stałe i ciecze mają ciężar właściwy znacznie większy od ciężaru właściwego powietrza, ciała te „toną" w powietrzu. Wszystkie ciała w naszym otoczeniu doznają ze strony powietrza parcia do góry. Wartość tego wyporu najczęściej pomijamy, chyba

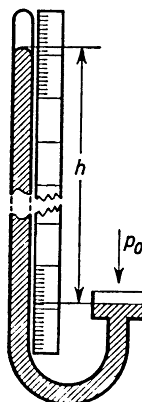
że robimy bardzo dokładne pomiary lub mamy do czynienia z ciałami o dużych objętościach. Balony wypełnione gazem lżejszym od powietrza (lub powietrzem ogrzanym) unoszą się do góry pod wpływem siły równej różnicy wyporu i ciężaru własnego balonu. Siła ta, zwana *siłą nośną*, powoduje ruch balonu do góry. Wartość siły nośnej podczas wznoszenia się balonu maleje z tego względu, że ciężar właściwy górnych warstw powietrza jest mniejszy niż warstw przy powierzchni Ziemi. Balon zatem może wznieść się do określonej wysokości i tam zawisnąć. Dalsze wznoszenie się można wywołać przez zmniejszanie ciężaru balonu, np. przez wyrzucanie balastu.

9.3. Ciśnienie atmosferyczne

Stan atmosfery ziemskiej można charakteryzować przez podawanie gęstości powietrza na różnych wysokościach, gdyż dzięki temu, że gazy są ściśliwe, ustala się pod wpływem siły ciężkości pewien rozkład gęstości powietrza. Ponieważ jednak łatwiej jest mierzyć ciśnienie powietrza niż jego gęstość, więc zazwyczaj mierzone są ciśnienia panujące na różnych wysokościach. Do pomiaru ciśnienia atmosferycznego stosuje się barometry. Działanie znanych z kursu szkolnego barometrów rtęciowych, których budowę przypominają rys. 9.15 i 9.16, opiera się na prawie hydrostatycznym równo-



Rys. 9.15



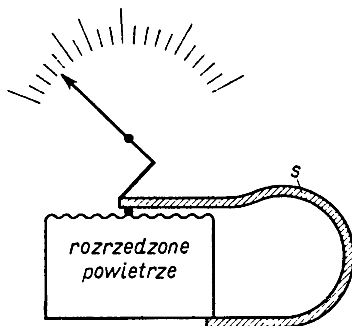
Rys. 9.16

wagi cieczy w naczyniach połączonych (por. p. 9.1.3). Istotną rzeczą przy pomiarze ciśnienia za pomocą barometru jest dokładne odczytanie wysokości słupa rtęci i ustalenie temperatury. Zmierzone ciśnienia powinny być tak przeliczone, by odpowiadały wykonaniu pomiaru w temperaturze 0°, na poziomie morza i na szerokości geograficznej 45°.

Istnieją specjalne tablice poprawek barometrycznych, uwzględniające zarówno zmianę gęstości rtęci z temperaturą, jak i rozszerzalność używanych w barometrach skal metalowych. Tablice te uwzględniają również różnice wartości przyspieszenia

ziemskiego w zależności od szerokości geograficznej i wzniesienia nad poziom morza. Wprowadzenie tych poprawek nosi nazwę *redukcji wskazań barometru do zera*.

Barometry rtęciowe są niewygodne w użyciu ze względu na swe rozmiary i stosowanie bądź co bądź kruchych rurek szklanych. Zastępujemy je często barometrami metalowymi, zwanymi *aneroidami*. Puszka metalowa, zaopatrzona w pofałdowaną szczelną pokrywkę (rys. 9.17), wypełniona jest rozrzedzonym powietrzem. Sprężyna S zapobiega wgnieceniu pokrywki do wnętrza na skutek działania ciśnienia atmosferycznego. Koniec sprężyny S połączony jest ze wskazówką za pomocą dźwigni. Zmiany ciśnienia zewnętrznego wywołują przesuwanie się wskazówki na tle podziałki.



Rys. 9.17

Aneroid musi być wykalibrowany według wskazań barometru rtęciowego. Ze względu na zmiany właściwości sprężystych metalu, zachodzące z biegiem czasu, wskazania aneroidu muszą być korygowane okresowo przez porównanie ze wskazaniami barometru rtęciowego.

Zakładając, że do powietrza można stosować prawo Boyle'a–Mariotte'a i że temperatura powietrza jest wszędzie jednakowa, równa 0°C, można wyprowadzić wzór na ciśnienie p_h panujące na wysokości h nad Ziemią:

$$p_h = p_0 e^{-\varrho_0 g h / p_0},$$

gdzie e = 2,718 ... oznacza podstawę logarytmów naturalnych, ϱ_0 — gęstość powietrza w tzw. warunkach normalnych, tzn. w temperaturze 0°C przy ciśnieniu normalnym $p_0 = 1{,}01325 \cdot 10^5$ N/m^2 (czyli 1 atm). Jest to jedna z postaci tzw. *wzoru barometrycznego*. Wzór ten jest oczywiście tylko w przybliżeniu słuszny, gdyż założenia, na jakich się opiera, nie są dokładnie spełnione.

Warunki panujące w atmosferze charakteryzuje tab. 9.2, podająca średnie wartości ciśnień, gęstości i temperatur dla różnych warstw atmosfery. Tabela ta dotyczy tzw. troposfery, sięgającej do 10–15 km i obejmującej 79% masy całego powietrza. Powyżej troposfery znajduje się stratosfera, sięgająca do 80–100 km i obejmująca 20% masy atmosfery. Ponad stratosferą rozciąga się tzw. jonosfera, sięgająca do 3200 km. Ciśnienie panujące na dolnej granicy jonosfery, na wysokości 100 km, wynosi ok. 0,75 N/m^2 (czyli ok. 0,006 mm Hg), a masa całej jonosfery nie przekracza 1% całej masy atmosfery.

Oczywiście im wyżej, tym ciśnienie staje się niższe, dążąc do zera na górnej granicy atmosfery ziemskiej.

Tabela 9.2

Średni stan atmosfery ziemskiej na różnych wysokościach

Wysokość nad poziomem morza m	Ciśnienie 10^5 N/m^2	Gęstość względna w stosunku do powietrza w warunkach normalnych	Temperatura °C
0	1	1	15
1 000	0,887	0,907	8,5
2 000	0,784	0,822	2
3 000	0,692	0,742	−4,5
4 000	0,608	0,669	−11
5 000	0,533	0,601	−17,3
10 000	0,261	0,337	−50

9.4. Manometry. Pompy rozrzedzające

Manometry są to urządzenia służące do pomiaru ciśnienia gazu. Rozróżniamy manometry wypełnione cieczami (przede wszystkim wodne i rtęciowe), manometry metalowe oraz manometry specjalne, tzw. próżniowe, do mierzenia ciśnień bardzo niskich.

Pytania. 1. Czy ciśnienie p gazu zawartego w zbiorniku przedstawionym na rys. 9.18 jest wyższe, czy niższe od ciśnienia atmosferycznego p_0? 2. Dlaczego stosunek wysokości h_w/h_r w manometrach wodnym i rtęciowym dołączonych do tego samego zbiornika z gazem (rys. 9.19) jest w przybliżeniu równy 13,6:1?

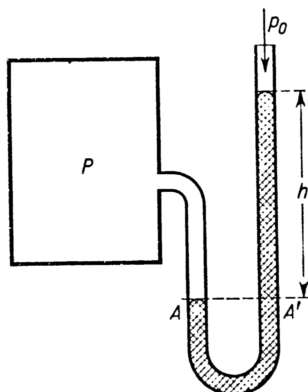
Ogólnie biorąc, ciecze manometryczne nie powinny być zbyt lepkie, jak gliceryna, oliwa; zbyt lotne — jak eter; nie powinny wchodzić w reakcje chemiczne z gazem, którego ciśnienie jest mierzone i nie powinny wykazywać dużej rozszerzalności cieplnej.

Działanie manometrów metalowych opiera się na właściwościach sprężystych ciał stałych.

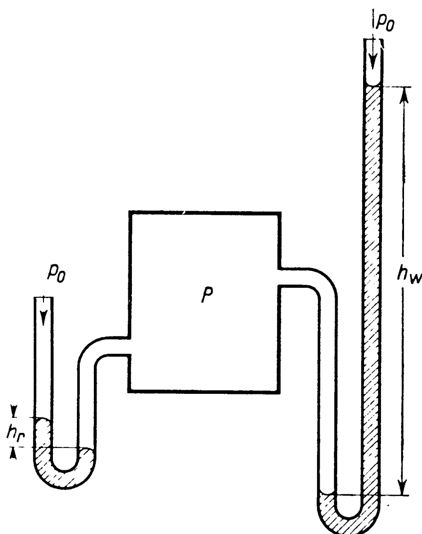
Wiele typów manometrów metalowych umożliwia automatyczną rejestrację wyników pomiarów. Swobodny koniec rurki odchylającej się mniej lub więcej, w zależności od mierzonego ciśnienia, połączony jest za pośrednictwem układu dźwigni z rysikiem. Koniec rysika, dotykający powierzchni papieru milimetrowego przewijanego ruchem jednostajnym z jednego bębna na drugi, kreśli na papierze przebieg zmian ciśnienia w określonym czasie.

Do mierzenia niskich ciśnień są m. in. szeroko stosowane manometry elektryczne, tzw. manometry Piraniego. Działanie takiego manometru opiera się na: 1) zależności strat ciepła ogrzewanego przewodnika (a więc i jego temperatury) od ciśnienia otaczającego gazu, 2) zależności oporu elektrycznego tego przewodnika od temperatury.

Praktycznie obie zależności wykorzystuje się w ten sposób, że drucik o dużej zależności oporu elektrycznego od temperatury umieszcza się w obszarze wysokiej próżni, gdzie ma być mierzone ciśnienie. Drucik stanowi jedną gałąź mostka Wheatstone'a (por. § 20.5) i jest grzany odpowiednim, stałym prądem elektrycznym. Im wyższa jest próżnia, tym bardziej utrudniona jest cyrkulacja gazu i tym mniejsze są straty cieplne. Przy stałym natężeniu prądu elektrycznego odpływ ciepła jest proporcjonalny do ciśnienia. A zatem pośrednio temperatura drucika zależy od ciśnienia. Mierząc opór drucika i znając zależność oporu od temperatury, można wyznaczyć temperaturę, a następnie w oparciu o znaną zależność temperatury od ciśnienia ustalić wartość ciśnienia. Zakres pomiaru ciśnienia za pomocą manometru Piraniego rozciąga się od 0,133–0,565 N/m² (czyli od 0,001 do 0,005 mm Hg).



Rys. 9.18



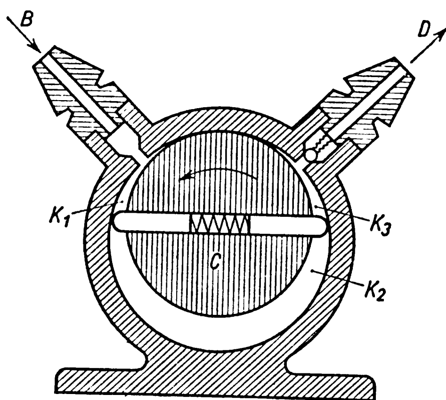
Rys. 9.19

Do pomiaru niskich ciśnień stosuje się dziś różne precyzyjne przyrządy. Ta dziedzina techniki pomiarowej rozwijała się szeroko w ostatnich dziesiątkach lat łącznie z techniką otrzymywania wysokiej próżni. Nowoczesne lampy rentgenowskie, lampy elektronowe wszelkich typów, aparatura fizyki jądrowej — oto kilka przykładów urządzeń wymagających wysokiej próżni. Przez osiągnięcie wysokiej próżni rozumiemy rozrzedzenie gazu do ciśnienia niższego od 0,1 N/m².

Jednym ze sposobów otrzymywania niskich ciśnień jest stosowanie *pomp rozrzedzających*. Omówimy działanie pomp rotacyjnych i dyfuzyjnych. Budowę *pompy rotacyjnej olejowej* przedstawia schematycznie rys. 9.20. W masywnym, metalowym cylindrze osadzony jest mimośrodowo walec metalowy *C*. Jak widać z rysunku, walec przylega szczelnie (poprzez warstwę oleju) w górnej części do ściany cylindra. Walec jest przecięty wzdłuż średnicy. W przecięciu osadzone są dwie łopatki przylegające również szczelnie do ścian cylindra dzięki rozsuwającemu działaniu silnej sprężyny.

Łopatki dzielą przestrzeń między powierzchnią cylindra i powierzchnią walca na trzy nie połączone ze sobą komory K_1, K_2 i K_3. Walec jest wprawiony w szybki ruch obrotowy (300–400 obrotów na minutę). Podczas ruchu walca objętości komór ulegają zmianie narastając stopniowo od zera do pewnej wartości maksymalnej, a potem znów malejąc do zera.

W cylindrze zewnętrznym mamy dwa wyloty. Wylot B służy do łączenia pompy ze zbiornikiem, z którego gaz ma być usunięty. Wylot D zaopatrzony w zawór pozwala na łączenie wnętrza pompy z otaczającą atmosferą.



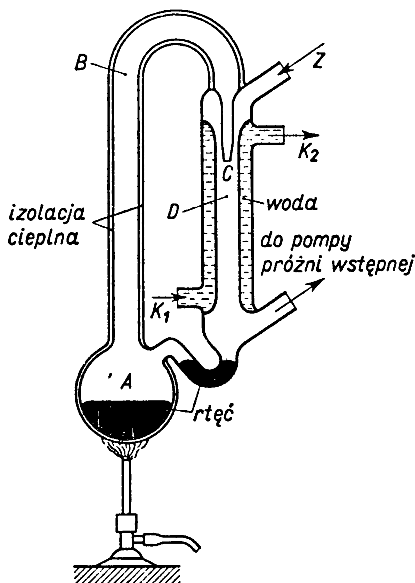
Rys. 9.20

Niech koniec łopatki mija w pewnej chwili wylot B podczas obrotu walca w kierunku zaznaczonym strzałką na rysunku. Stopniowo rośnie objętość komory K_1, przy czym gaz ze zbiornika jest zasysany do pompy. Równocześnie podczas ruchu walca maleje objętość komory K_3. W poprzednich stadiach ruchu komory K_3 i K_2 były w takim położeniu, w jakim obecnie jest K_1, tzn. w poprzednich fazach ruchu gaz ze zbiornika dochodził do komory K_3, a następnie K_2. Podczas zmniejszenia się objętości komory K_3 rośnie prężność zawartego w niej gazu osiągając wreszcie taką wartość, że otwarty zostaje zawór wylotu D umożliwiając ujście gazu do atmosfery. Ostatecznie gaz z komory K_3 zostaje wyparty, a sama komora K_3 przestaje istnieć na krótki czas, aby wystąpić znów przy dalszym obrocie walca, ale już w charakterze komory K_1. W tym czasie objętość komory K_2, przekształconej na K_3, stopniowo się zmniejsza i znów gaz swoim ciśnieniem otwiera zawór wylotowy i przedostaje się do otaczającej atmosfery. Przy każdym obrocie walca powtarza się omówione wyżej zasysanie gazu ze zbiornika i usuwanie go na zewnątrz.

W nowoczesnych pompach rotacyjnych olejowych cała opisana część robocza jest umieszczona w dużym zbiorniku z olejem. Olej w pompach rotacyjnych spełnia podwójną rolę: służy jako smar, a równocześnie działa uszczelniająco. Za pomocą pojedynczej pompy rotacyjnej można osiągnąć rozrzedzenie (czyli, jak się często mówi, próżnię) o ciśnieniu rzędu kilku N/m^2. Połączenie szeregowe 2–3 takich pomp może dać próżnię rzędu tysięcznych części N/m^2.

(Może warto tu wspomnieć, że po dołączeniu wylotu D do zbiornika, a wylotu B do atmosfery, przy pozostałych nie zmienionych warunkach pracy, gaz jest zasysany z atmosfery i wtłaczany do zbiornika. W tym przypadku pompa rotacyjna działa jako pompa zgęszczająca, czyli kompresor. Osiągane ciśnienia mogą dochodzić do $6 \cdot 10^5$ N/m².)

Pompa dyfuzyjna rtęciowa jest przedstawiona na rys. 9.21. Zbiornik A zawiera rtęć ogrzewaną do wrzenia. Pary rtęci przechodzą do góry rurką B odizolowaną cieplnie od otoczenia. Przez zwężony wylot C para rtęci przedostaje się do naczynia D chłodzonego od zewnątrz strumieniem zimnej wody płynącej od K_1 do K_2. W tej części pompy para rtęci ulega skraplaniu i spływa z powrotem do zbiornika A.



Rys. 9.21

W ścianach naczyń C i D mamy dwa odprowadzenia: górne łączy pompę ze zbiornikiem Z, w którym ma być wytworzona wysoka próżnia, dolne zaś daje połączenie z przestrzenią, w której jest utrzymywana tzw. próżnia wstępna. Próżnia wstępna jest wytworzona wskutek działania pompy pomocniczej, np. olejowej rotacyjnej. Strumień pary rtęci opadający ku dołowi porywa ze sobą cząstki gazu ze zbiornika (do istoty tego zjawiska wrócimy jeszcze w p. 9.6.1). W dolnej części naczynia D cząstki te są usuwane wskutek działania pompy pomocniczej wytwarzającej próżnię wstępną. Prężność pary nasyconej rtęci stanowi granicę, do której obniżyć można ciśnienie w zbiorniku. Ciśnienie pary nasyconej rtęci w temperaturze 20°C wynosi ok. 0,174 N/m² (czyli 0,00131 mm Hg). Jest ono bardzo niewielkie w porównaniu z ciśnieniem atmosferycznym i można je zaniedbywać, mówiąc np. o „próżni Torricellego" w obszarze nad rtęcią w barometrze rtęciowym. Jest ono jednak duże, gdy chodzi o uzyskanie

wysokiej próżni. Aby zmniejszyć wpływ tego ciśnienia, stosuje się specjalne chłodzenie części aparatury łączącej naczynie D ze zbiornikiem z gazem rozrzedzanym. Do chłodzenia często używane jest ciekłe powietrze obniżające temperaturę do około —180°C. Jest to tzw. wymrażanie par rtęci obniżające bardzo znacznie prężność pary. W ten sposób można uzyskać próżnię rzędu 0,0013 N/m^2 (czyli 10^{-5} mm Hg).

Obok pomp dyfuzyjnych rtęciowych są też w użyciu *pompy dyfuzyjne olejowe*, w których rtęć zastąpiona jest olejem o małej prężności pary nawet w temperaturze pokojowej. Są to zazwyczaj oleje apiezonowe lub silikonowe. Ciśnienie pary nasyconej tych olejów w temperaturze pokojowej wynosi około 0,000 010 N/m^2 (czyli ok. 10^{-7} mm Hg). Użycie ich umożliwia otrzymywanie w pompie dyfuzyjnej próżni rzędu 10^{-5} N/m^2.

Pompy dyfuzyjne olejowe i rtęciowe działają tylko wtedy, gdy próżnia wstępna jest rzędu co najmniej 10^1 N/m^2 (dziesiątych części milimetra Hg). Im lepsza próżnia wstępna, tym szybsze działanie pompy dyfuzyjnej.

Bardzo niskie ciśnienia można też otrzymywać wykorzystując zjawisko adsorpcji gazów na powierzchniach metali lub węgla aktywowanego. W ten sposób osiągnięto w niewielkich objętościach próżnię rzędu 10^{-7} N/m^2. Celem zwiększenia adsorpcji gazów na powierzchni węgla aktywowanego oziębia się go do temperatury ciekłego powietrza.

Z obliczeń wynika, że najmniejsze ciśnienia panują w przestrzeniach międzygalaktycznych i wynoszą około 10^{-14} N/m^2. Stopień rozrzedzenia jest tam tak wielki, że przeciętnie 1 atom przypada na objętość 1 dm^3. W warunkach normalnych gazu średnio na 1 dm^3 przypada $3 \cdot 10^{22}$ cząsteczek.

Największe ciśnienia panują w gwiazdach zwanych białymi karłami. Gęstość ich jest szacowana na dziesiątki milionów kilogramów na metr sześcienny.

9.5. Ogólna charakterystyka ruchu płynów

Rozważania hydro- i aerodynamiczne mają podstawowe znaczenie w technice, m. in. wszędzie tam, gdzie wchodzi w grę przepływ cieczy lub gazu przez przewody (wodociągi, naftociągi, przewody gazu świetlnego, gazu ziemnego itp.). Hydrodynamika i aerodynamika zajmują się także bardzo ważnymi dla współczesnej komunikacji zjawiskami towarzyszącymi ruchowi ciał stałych w ośrodkach ciekłych i gazowych, sprowadzając zresztą zwykle te zagadnienia do zagadnień opływania nieruchomego ciała przez strumień płynu (badania eksperymentalne prowadzone w tunelach aerodynamicznych).

Ścisłe matematyczne rozwiązanie wspomnianych zagadnień napotyka częstokroć trudności nie do przezwyciężenia, nic więc dziwnego, że i w tej dziedzinie stosuje się różne „modele". Dzięki temu ujęcie matematyczne staje się znacznie prostsze, a wyniki — mimo że przybliżone — mają jednak znaczenie praktyczne. Większość naszych dalszych rozważań odnosić się będzie do modelowych płynów, zwanych doskonałymi, i modelowych, wyidealizowanych rodzajów ruchów (stacjonarnych, laminarnych).

Pod nazwą płynów doskonałych rozumieć będziemy ciecze i gazy nieściśliwe i nielepkie*.

O ściśliwości cieczy i gazów już mówiliśmy. Założenie braku ściśliwości cieczy nie budzi żadnych zastrzeżeń, natomiast w pierwszej chwili może się wydawać, że założenie braku ściśliwości gazów nie ma realnych podstaw. Okazuje się jednak, że zmiany ciśnień powstające w gazie podczas jego ruchu z prędkościami mniejszymi od prędkości głosu są zazwyczaj tak małe, że towarzyszące im zmiany objętości i gęstości można w przybliżeniu zaniedbywać. O lepkości (czyli tarciu wewnętrznym) będziemy szerzej mówili w § 9.7 i p. 16.11.3.

Przejdźmy do dalszych uwag ogólnych, dotyczących ruchu płynów. Chcemy od razu na wstępie podkreślić, że opis ruchu płynu będzie na ogół odmienny od tego, który stosowaliśmy w kinematyce punktu materialnego. Najczęściej nasza uwaga nie będzie się koncentrowała na poszczególnych cząstkach płynu: nie będziemy mówili o tym, jak z biegiem czasu zmieniają się współrzędne (x, y, z) określające ich położenie, jak zmienia się ich prędkość i przyspieszenie. Będziemy raczej brali pod uwagę poszczególne punkty przestrzeni z obszaru, przez który przepływa płyn i będziemy ustalali w każdym z tych punktów prędkość przepływu $v(x, y, z)$, gęstość płynu $\varrho(x, y, z)$ oraz ciśnienie $p(x, y, z)$ w funkcji czasu.

Badania wykazały, że przepływ płynu może mieć bardzo różny charakter. Podamy zatem kilka uwag klasyfikacyjnych.

Przepływ nazywamy *stacjonarnym* (w odróżnieniu od niestacjonarnego), gdy w określonym punkcie przestrzeni prędkość $\mathbf{v}$ przepływu płynu jest stała, niezależna od czasu. W sąsiednim punkcie prędkość może być inna, lecz również jest niezależna od czasu.

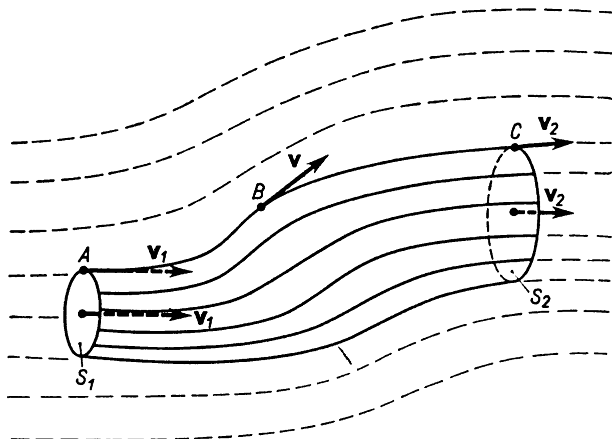
Przepływ nazywamy *stacjonarnym warstwowym* lub krótko *laminarnym* (w odróżnieniu od przepływu burzliwego), gdy wszystkie cząstki płynu poruszają się po torach równoległych do siebie. Wtedy ruch płynu sprowadza się do przesuwania się (poślizgu) warstw płynu względem siebie. Przepływ laminarny ma charakter najprostszy; jemu też poświęcimy dalej najwięcej uwagi.

W ilościowym ujęciu zagadnień związanych z przepływem płynu przydatne jest pojęcie strumienia masy lub strumienia objętości płynu. Zajmiemy się nieco szerzej tą nową wielkością fizyczną, gdyż ma ona ogólniejsze znaczenie. Mogliśmy np. już wcześniej wprowadzić ten termin mówiąc o mocy. Aczkolwiek terminologia nie jest ściśle ujednolicona, to jednak w wielu przypadkach *strumień wielkości fizycznej X* (np. *masy*, *objętości*, *energii*), odpowiadający pewnej *powierzchni*, oznacza pochodną tej wielkości względem czasu, czyli dX/dt. Tak więc strumieniem energii, odpowiadającym pewnej powierzchni, jest dE/dt, tzn. ilość energii przenikającej w jednostce czasu przez badaną powierzchnię. W interesującym nas obecnie zagadnieniu *strumieniem masy płynu przepływającego przez daną powierzchnię* jest dm/dt, a *strumieniem objętości* dV/dt.

* W rozważaniach obecnych interesują nas te dwa warunki narzucane cieczom i gazom doskonałym. Dla uniknięcia nieporozumień trzeba jednak podkreślić, że w innych zagadnieniach inaczej definiujemy doskonałość gazu (por. § 9.2).

Ze strumieniem określonej wielkości ściśle się kojarzy nowa wielkość, a mianowicie gęstość strumienia. *Gęstość strumienia wielkości X* znajdujemy jako stosunek strumienia wielkości X przechodzącego prostopadle przez powierzchnię S do tej powierzchni. Tak więc gęstością strumienia energii (masy, objętości itp.) jest energia (masa, objętość itp.) przenikająca w jednostce czasu przez jednostkę powierzchni ustawionej prostopadle względem kierunku przepływu energii (masy, objętości itp.).

Wróćmy jednak do zagadnień interesujących nas obecnie i spróbujmy przez wprowadzenie takich pojęć, jak linie i rurki prądu dojść do nowej interpretacji strumienia masy i objętości oraz gęstości strumienia masy i objętości.



Rys. 9.22

W obszarze przepływu płynu weźmy pod uwagę niewielką powierzchnię S_1 (rys. 9.22) ustawioną prostopadle do kierunku przepływu płynu. Wobec założonych małych rozmiarów tej powierzchni można przyjąć, że prędkość $\mathbf{v}_1$ przepływu płynu jest w każdym punkcie tej powierzchni taka sama, prostopadła do S_1. Jeżeli założymy stacjonarność ruchu, to każda z kolejnych cząstek płynu, dopływających do dowolnego punktu A, leżącego na konturze powierzchni S_1, poruszać się będzie po identycznym torze w przestrzeni. Tor ten będzie miał taki przebieg, że wektor prędkości $\mathbf{v}$ (zmienny wzdłuż toru, lecz stały dla każdego określonego jego punktu) będzie stale do niego styczny. Odcinek jednego z takich torów przedstawia linia ABC na rys. 9.22. Tak zdefiniowane tory nazywamy *liniami prądu.*

Wyobraźmy sobie, że analogiczne linie prądu nakreśliliśmy dla wszystkich punktów stanowiących kontur powierzchni S_1 i otrzymaną w ten sposób tzw. rurkę prądu zamknęliśmy przekrojem S_2 również prostopadłym do prędkości przepływu $\mathbf{v}_2$ charakterystycznej dla całego przekroju S_2.

Rurka prądu spełnia następujące warunki: 1) przez powierzchnię boczną rurki utworzoną przez linie prądu nie odbywa się przepływ płynu, 2) płyn wpływa do rurki wyłącznie przez jeden jej koniec (np. przez przekrój S_1), wypływa — przez drugi (np. przez S_2).

Załóżmy, że wewnątrz rurki nie ma ani źródeł płynu, ani jego zaniku (upływu). W tych warunkach strumień masy wpływający przez S_1 równa się strumieniowi masy wypływającemu przez S_2, czyli

$$\varrho_1 v_1 S_1 = \varrho_2 v_2 S_2, \tag{9.6}$$

gdzie ϱ_1 i ϱ_2 oznaczają gęstości płynu odpowiednio na przekrojach S_1 i S_2. Jest to tzw. *równanie ciągłości strumienia masy.*

Gęstość strumienia masy na każdym z badanych przekrojów wynosi odpowiednio $\varrho_1 v_1$ i $\varrho_2 v_2$. Łatwo sprawdzić, że gęstości strumienia masy na obu przekrojach rurki prądu są odwrotnie proporcjonalne do powierzchni tych przekrojów:

$$\frac{\varrho_1 v_1}{\varrho_2 v_2} = \frac{S_2}{S_1}.$$

W odniesieniu do płynu nieściśliwego (np. cieczy doskonałej) równanie ciągłości strumienia masy (9.6) upraszcza się (dzięki temu, że $\varrho_1 = \varrho_2$) i przechodzi w *równanie ciągłości strumienia objętości*

$$v_1 S_1 = v_2 S_2. \tag{9.7}$$

Gęstość strumienia objętości na obu przekrojach liczbowo odpowiada wartościom prędkości v_1 i v_2. Z (9.7) wynika, że gęstości strumieni objętości są również odwrotnie proporcjonalne do odpowiednich przekrojów.

$$\frac{v_1}{v_2} = \frac{S_2}{S_1}. \tag{9.8}$$

Treść równania (9.8) można też ująć w ten sposób, że im węższy jest przekrój *rurki prądu, tym większa jest prędkość przepływu płynu.*

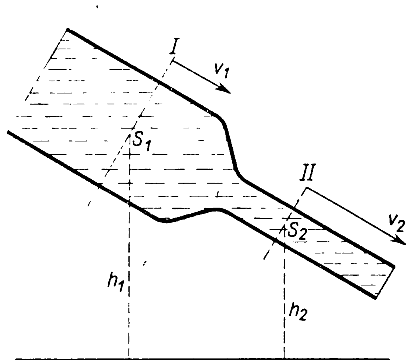
Obrazowe przedstawienie omówionych zależności otrzymamy po wprowadzeniu umowy, że linie prądu nie będą rysowane dowolnie gęsto, lecz z zachowaniem takich proporcji, by *liczba linii prądu przebijających prostopadle jednostkę powierzchni była równa gęstości strumienia objętości* (czyli liczbowo równa v). Przy stosowaniu takiej umowy z gęstości przebiegu linii prądu w różnych obszarach przestrzeni, w których odbywa się przepływ płynu, można sądzić bezpośrednio o prędkości przepływu. *Im większa jest gęstość linii, tym większa prędkość przepływu płynu w danym obszarze.*

Warto może podkreślić, że wprowadzenie na pozór abstrakcyjnego pojęcia rurek prądu ma duże znaczenie praktyczne ze względu na to, że ściany rzeczywistych przewodów, rur, przez które przepływają płyny, stwarzają ograniczenia boczne prądów analogiczne do istniejących w rurkach prądu. A zatem wnioski wyprowadzone dla rurek prądu można stosować do przepływu płynu przez rzeczywiste przewody, oczywiście przy spełnieniu odpowiednich założeń co do rodzaju płynu i charakteru przepływu. Wiemy już wobec tego, jak zmieniają się prędkości w zależności od przekroju przewodów przy przepływie laminarnym nieściśliwego płynu. W następnym punkcie dowiemy się, jakie zależności obowiązują w odniesieniu do ciśnień na różnych przekrojach.

9.6. Prawa rządzące ruchem cieczy doskonałej

9.6.1. *Prawo Bernoulliego.* Prawo Bernoulliego dotyczy przepływu cieczy doskonałej przez przewody o zmiennym przekroju. Wiąże ono ciśnienia i prędkości przepływu na poszczególnych przekrojach z powierzchniami tych przekrojów i ich wysokościami względem obranego poziomu odniesienia. Wyprowadzenie tego prawa opiera się na rozważaniach energetycznych.

Rozpatrzmy przepływ cieczy doskonałej przez przewód o niejednakowym przekroju (rys. 9.23), ustawiony ukośnie względem poziomu. Niech I i II przedstawiają dwa przekroje poprzeczne o polach S_1 i S_2, znajdujące się odpowiednio na wysokościach h_1 i h_2 nad dowolnie wybranym poziomem odniesienia. Ciśnienia na tych przekrojach niech wynoszą odpowiednio p_1 i p_2, a prędkości przepływu $\mathbf{v}_1$ i $\mathbf{v}_2$.



Rys. 9.23

Kosztem różnicy pracy W_1, włożonej w układ w krótkim czasie Δt na przekroju I, i pracy W_2, wykonanej w tym czasie na przekroju II, oraz kosztem różnicy energii potencjalnych grawitacyjnych E_{p1} i E_{p2} mas cieczy przepływających w czasie Δt przez oba przekroje — powstaje różnica energii kinetycznych wymienionych mas cieczy. A zatem

$$W_1 - W_2 + E_{p1} - E_{p2} = E_{k2} - E_{k1}. \tag{9.9}$$

Wymienione prace można przedstawić w postaci

$$\begin{aligned} W_1 &= p_1 S_1 v_1 \Delta t, \\ W_2 &= p_2 S_2 v_2 \Delta t, \end{aligned} \tag{9.10}$$

gdyż przesunięcia w krótkim czasie Δt równają się odpowiednio $v_1 \Delta t$ i $v_2 \Delta t$.

Wymienione wyżej energie potencjalne można przedstawić jako

$$\begin{aligned} E_{p1} &= mgh_1 = S_1 v_1 \Delta t \varrho g h_1, \\ E_{p2} &= mgh_2 = S_2 v_2 \Delta t \varrho g h_2. \end{aligned} \tag{9.11}$$

Podobnie wymienione wyżej energie kinetyczne można przedstawić jako

$$E_{k1} = \frac{mv_1^2}{2} = \frac{\varrho S_1 v_1 \Delta t v_1^2}{2},$$
$$E_{k2} = \frac{mv_2^2}{2} = \frac{\varrho S_2 v_2 \Delta t v_2^2}{2}. \tag{9.12}$$

Przy wypisywaniu (9.11) i (9.12) uwzględniliśmy niezmienną gęstość cieczy doskonałej w całej objętości ($\varrho_1 = \varrho_2 = \varrho$).

Podstawiając (9.10), (9.11) i (9.12) do (9.9) otrzymujemy

$$p_1 S_1 v_1 \Delta t - p_2 S_2 v_2 \Delta t + S_1 v_1 \Delta t \varrho g h_1 - S_2 v_2 \Delta t \varrho g h_2 = \tfrac{1}{2} S_2 v_2 \Delta t \varrho v_2^2 - \tfrac{1}{2} S_1 v_1 \Delta t \varrho v_1^2.$$

Uwzględniając równanie ciągłości strumienia objętości dochodzimy do wyrażenia

$$p_1 - p_2 + \varrho g h_1 - \varrho g h_2 = \tfrac{1}{2} \varrho v_2^2 - \tfrac{1}{2} \varrho v_1^2,$$

skąd

$$p_1 + \varrho g h_1 + \tfrac{1}{2} \varrho v_1^2 = p_2 + \varrho g h_2 + \tfrac{1}{2} \varrho v_2^2.$$

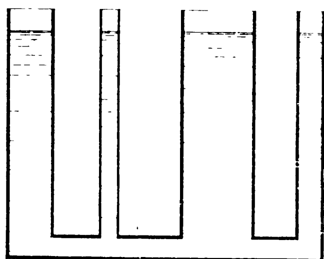
Przekroje *I* i *II* były wybrane zupełnie dowolnie, a więc ogólnie można napisać

$$p + \varrho g h + \tfrac{1}{2} \varrho v^2 = \text{const}. \tag{9.13}$$

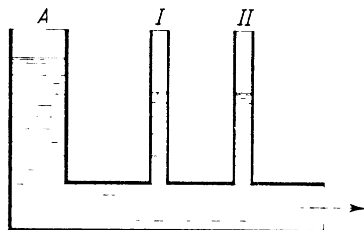
Równanie to wyraża *treść prawa Bernoulliego* w ujęciu ogólnym.

W przypadku szczególnym przepływu cieczy doskonałej przez przewód poziomy nie ma zmiany energii potencjalnej ($h_1 = h_2 = h = \text{const}$), a zatem (9.13) redukuje się do postaci

$$p + \tfrac{1}{2} \varrho v^2 = \text{const}. \tag{9.14}$$



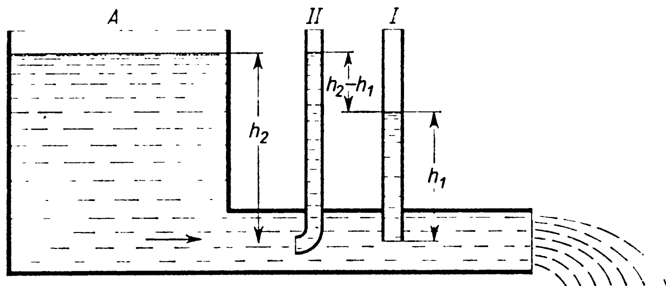
Rys. 9.24



Rys. 9.25

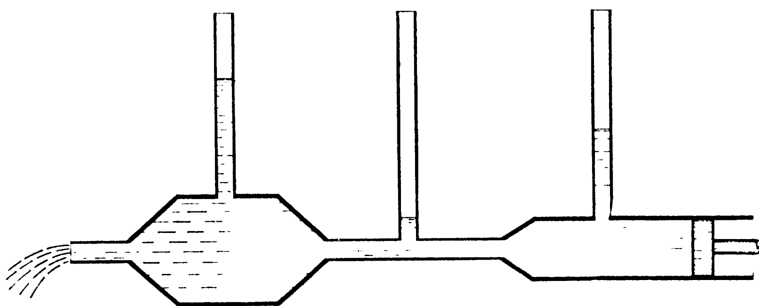
Omówimy kilka wniosków wypływających z prawa Bernoulliego. Wiemy, że w hydrostatyce obowiązuje prawo, że na określonym poziomie w danej cieczy panuje wszędzie to samo ciśnienie. Równowaga cieczy w naczyniach połączonych zachodzi wtedy, gdy we wszystkich ramionach, niezależnie od ich przekrojów, ciecz sięga do tego samego poziomu (rys. 9.24). Inaczej przedstawia się sprawa rozkładu ciśnień podczas przepływu cieczy doskonałej przez przewody poziome. W czasie przepływu cieczy doskonałej przez przewód o jednakowym przekroju (rys. 9.25) w rurkach manometrycznych *I* i *II* poziomy są jednakowe, lecz niższe niż w naczyniu *A*. Jeżeli jednak jednej

z rurek nadamy kształt wygięty (rurka *II* na rys. 9.26) i zwrócimy ją płaszczyzną wylotu w stronę strumienia napływającej cieczy (tzw. *rurka Pitot*), to sytuacja się zmieni: w rurce *II* ustali się poziom taki jak w naczyniu *A*, a w rurce *I* — tak jak poprzednio — poziom niższy. Ciśnienie hydrostatyczne cieczy w rurce *II* jest miarą pełnego ciśnienia, pod którym odbywa się wypływ, zwanego *ciśnieniem hydraulicznym*, równego $(p+\varrho v^2/2)$. Ciśnienie hydrostatyczne cieczy w rurce *I* jest miarą *ciśnienia p*, zwanego *piezometrycznym*, a ciśnienie hydrostatyczne słupa cieczy o wysokości (h_2-h_1) odpowiada *ciśnieniu hydrodynamicznemu* $\varrho v^2/2$, wynikającemu z ruchu cieczy



Rys. 9.26

Podczas przepływu cieczy doskonałej przez przewody o niejednakowym przekroju sprawa jeszcze się komplikuje. W przewężeniach (rys. 9.27), zgodnie z prawem Bernoulliego, mamy dużą prędkość przepływu, a niskie ciśnienie. W szerokich częściach przewodu małym prędkościom przepływu odpowiadają wysokie ciśnienia. Zaznacza się to w postaci słupów cieczy o różnych wysokościach w rurkach pionowych dołączonych



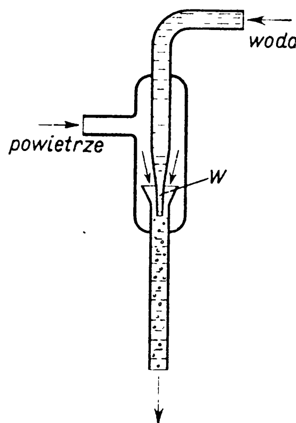
Rys. 9.27

do głównego przewodu. Trzeba jednak pamiętać, że wysokości tych słupów cieczy nie są miarą pełnego ciśnienia hydraulicznego przepływającej cieczy. Wyrażają one ciśnienia odpowiadające pierwszemu składnikowi sumy z równania Bernoulliego dla rur poziomych, czyli ciśnienia piezometryczne.

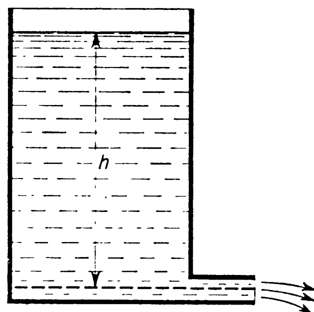
Zmniejszenie ciśnienia podczas przepływu cieczy przez przewężenia znalazło zastosowanie m. in. *w pompce wodnej* (rys. 9.28). Podczas przepływu strumienia wody

przez zwężony wylot rurki *W* występuje znaczna zniżka ciśnienia. Dzięki temu gaz z otoczenia (z rurki łączącej pompę ze zbiornikiem gazu) zostaje zasysany przez strumień cieczy. Pęcherzyki gazu odpływają razem ze strumieniem cieczy. W ten sposób w zbiornikach połączonych z pompką wodną można otrzymać zniżkę ciśnienia do ok. 1400 N/m^2 (ok. 10 mmHg).

9.6.2. *Prawo Torricellego.* Prawo Torricellego dotyczy prędkości wypływu cieczy doskonałej pod wpływem siły ciężkości z otworu znajdującego się na głębokości h w stosunku do powierzchni swobodnej cieczy w ścianie bocznej lub w dnie naczynia (rys. 9.29). Zaniedbujemy tarcie o ścianki naczynia i opór otaczającego ośrodka oraz zakładamy, że pole przekroju otworu jest bardzo małe w stosunku do pola powierzchni swobodnej cieczy.



Rys. 9.28



Rys. 9.29

Obserwując całość zjawiska w ciągu krótkiego czasu Δt stwierdzamy ubytek pewnej masy cieczy z poziomu pierwotnej powierzchni swobodnej (praktycznie niezmiennego wobec wprowadzonego założenia), a równocześnie wypływ takiej samej masy z prędkością v na poziomie o h niższym. Można zatem przyjąć, że kosztem ubytku energii potencjalnej grawitacyjnej, równego mgh, powstaje energia kinetyczna $\frac{1}{2}mv^2$.

$$mgh = \frac{1}{2}mv^2,$$

skąd

$$v = \sqrt{2gh}. \tag{9.15}$$

Wzór otrzymany wyraża treść *prawa Torricellego*: *prędkość wypływu cieczy doskonałej z niewielkiego otworu znajdującego się na głębokości h pod powierzchnią swobodną cieczy jest taka sama, jak prędkość ciała swobodnie spadającego w próżni z wysokości h.* Jak widać, prędkość wypływu cieczy doskonałej nie zależy od jej gęstości.

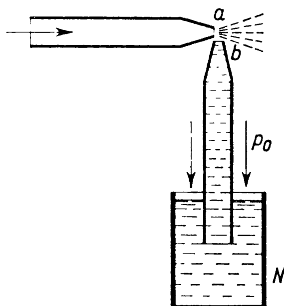
Przy długotrwałym wypływie cieczy prędkość wypływu stopniowo się zmniejsza, gdyż wysokość h maleje.

P y t a n i e. Prawo Torricellego można traktować jako wniosek z prawa Bernoulliego. W jaki sposób można to wykazać?

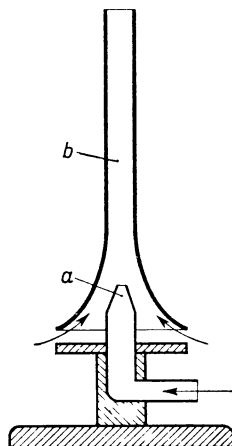
9.7. Ogólne uwagi dotyczące ruchu gazów

Wiemy już, że podstawowym prawem określającym ruch cieczy jest prawo Bernoulliego, odnoszące się do cieczy doskonałej. Wróćmy raz jeszcze do pytania, czy popełniamy wielki błąd stosując to prawo do ruchu gazów, tzn. zaniedbując ich ściśliwość.

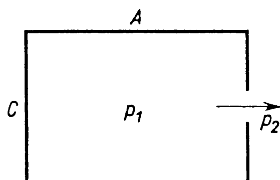
Z obliczeń wynika, że względne zmiany objętości $\Delta V/V$, występujące w gazie podczas jego ruchu o prędkości rzędu 100 m/s, czyli 360 km/h, są stosunkowo niewielkie — wynoszą około 6%. Tak więc nawet przy takich prędkościach, jak średnie prędkości samolotów, popełniamy stosunkowo niewielki błąd stosując równanie Bernoulliego do gazów bez wprowadzania dodatkowych poprawek, uwzględniających ich ściśliwość.



Rys. 9.30



Rys. 9.31



Rys. 9.32

Do przepływu gazów przez przewody o różnych przekrojach można zatem stosować wnioski wynikające z prawa Bernoulliego. W przewężeniach przewodów albo u zwężonych wylotów mamy większą prędkość przepływów, a zmniejszone ciśnienie. To tłumaczy np. działanie zasysające rozpylacza (rys. 9.30). U zwężonego wylotu rurki podczas przedmuchiwania strumienia gazu powstaje zniżka ciśnienia. Dzięki temu ciśnienie atmosferyczne p_0 wpycha ciecz z naczynia *N* do rurki *b*. Po osiągnięciu wysokości wylotu ciecz zostaje rozpylona wskutek działania strumienia przepływającego powietrza.

Rysunek 9.31 przedstawia zasadę działania palnika gazowego Bunsena. Strumień gazu świetlnego przechodzi przez przewężenie a i wytwarza tam zniżkę ciśnienia. Dzięki temu zasysane jest z zewnątrz powietrze. W przewodzie b następuje dokładne wymieszanie gazu i powietrza. Mieszanina ta dochodzi do zewnętrznego wylotu palnika, gdzie ulega spalaniu. Regulowanie płomienia polega na utrudnieniu lub ułatwieniu dopływu większej ilości powietrza.

Z prawa Bernoulliego wynika prawo Bunsena, określające prędkość wypływu gazów z małych otworów w zbiornikach. Niech w zbiorniku A (rys. 9.32) panuje ciśnienie p_1, a na zewnątrz — ciśnienie p_2, przy czym $p_1 > p_2$. Zakładamy, że wypływowi podlegają tylko cząstki gazu przy wylocie. Wewnątrz naczynia, np. przy ściance C, przepływu nie ma, czyli tam $v = 0$. Wobec tego stosując prawo Bernoulliego otrzymamy

$$p_1 = p_2 + \frac{\varrho v^2}{2},$$

skąd

$$v = \sqrt{\frac{2(p_1 - p_2)}{\varrho}}.$$

Prędkość wypływu gazu jest proporcjonalna do pierwiastka kwadratowego z nadwyżki ciśnienia w naczyniu ponad ciśnienie otoczenia i odwrotnie proporcjonalna do pierwiastka kwadratowego z gęstości. Jest to treść *prawa Bunsena*, które, jak widać, jest również wnioskiem z prawa Bernoulliego.

Z ostatniego wzoru wynika, że prędkości wypływu dwóch gazów w tych samych warunkach są odwrotnie proporcjonalne do pierwiastków kwadratowych z ich gęstości

$$\frac{v_1}{v_2} = \sqrt{\frac{\varrho_2}{\varrho_1}}.$$

Wzór ten można stosować do obliczania gęstości gazów. Ponieważ jednak prędkość wypływu jest wielkością raczej niewygodną do mierzenia, więc pomiar gęstości sprowadzamy do mierzenia czasów wypływu jednakowych objętości dwóch gazów. Objętość V_1 gazu pierwszego, wypływającego z prędkością v_1 z wylotu o powierzchni S w czasie t_1, wynosi

$$V_1 = v_1 S t_1.$$

Odpowiednio

$$V_2 = v_2 S t_2.$$

Z równości obu tych objętości wynika

$$\sqrt{\frac{2(p_1 - p_2)}{\varrho_1}}\, S t_1 = \sqrt{\frac{2(p_1 - p_2)}{\varrho_2}}\, S t_2,$$

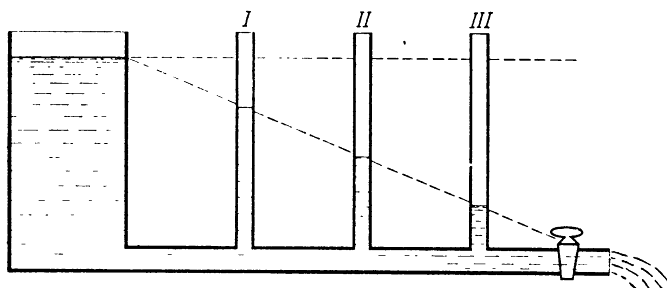
skąd

$$\frac{\varrho_1}{\varrho_2} = \frac{t_1^2}{t_2^2}.$$

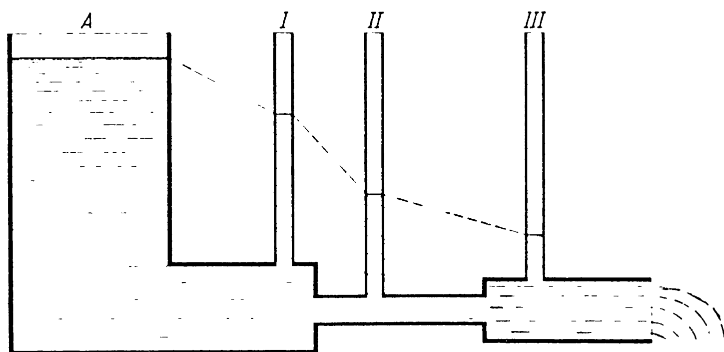
Stosunek gęstości dwóch gazów równa się stosunkowi kwadratów czasów potrzebnych do otrzymania jednakowych objętości wypływających gazów przy tej samej nadwyżce ciśnienia i tej samej powierzchni wylotu.

9.8. Lepkość płynów

Istnienie lepkości (zwanej także tarciem wewnętrznym) i jej wpływ na przepływ cieczy przez przewody widać wyraźnie z porównania rys. 9.25 i 9.33. Rys. 9.25 przedstawia teoretyczny rozkład ciśnień w cieczy doskonałej podczas jej przepływu przez przewód o niezmiennym przekroju, wynikający z prawa Bernoulliego. Rys. 9.33 natomiast przedstawia faktyczny rozkład ciśnień w analogicznym przewodzie podczas przepływu cieczy rzeczywistej. Widać równomierny spadek ciśnienia wzdłuż długości przewodu. Bardziej skomplikowany rozkład ciśnień w przewodzie o zmiennym przekroju przedstawia rys. 9.34.



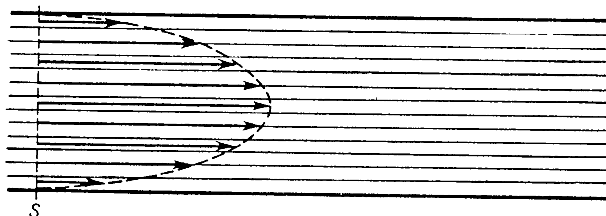
Rys. 9.33



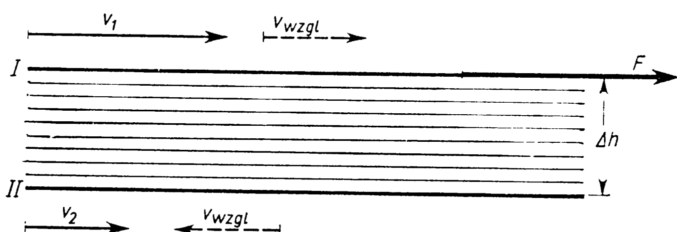
Rys. 9.34

Spadek ciśnienia wzdłuż drogi prądu cieczy wskazuje na istnienie siły oporu przeciwstawiającej się ruchowi, a więc pod tym względem podobnej do omawianego już tarcia. Przyjmuje się zazwyczaj, że ciecz przepływająca przez przewody przylega do ścian przewodu tworząc nieruchomą warstwę, coś w rodzaju cylindra czy kanału z cieczy, wewnątrz którego odbywa się ruch reszty cieczy.

W miarę przesuwania się od ścian przewodu do środka prędkość rośnie (rys. 9.35). Tak więc podczas przepływu cieczy mamy do czynienia z przesuwaniem się jednych jej warstw względem drugich i opór, jaki temu towarzyszy, słusznie może być nazywany tarciem wewnętrznym. W ogólnym użyciu jest jednak raczej krótsza nazwa — lepkość.



Rys. 9.35



Rys. 9.36

Właściwości różnych cieczy (a w ogólniejszym ujęciu — płynów) z punktu widzenia ich lepkości charakteryzuje wielkość zwana *współczynnikiem lepkości*. Definicja tego współczynnika opiera się na wynikach badań eksperymentalnych Newtona. Wyobraźmy sobie przepływ laminarny cieczy przedstawiony na rys. 9.36. Wyodrębniamy dwie warstwy I i II o powierzchni S, odległe od siebie o Δh. Warstwa górna porusza się z prędkością v_1, dolna — z prędkością v_2. Stałość tych prędkości jest zapewniona wtedy, gdy na górną warstwę cieczy działa siła zewnętrzna F, styczna do powierzchni, równoważąca opór ze strony cieczy. Newton wykazał, że siła ta wyraża się wzorem

$$F = \eta S \frac{\Delta v}{\Delta h}, \tag{9.16}$$

gdzie η oznacza współczynnik lepkości charakterystyczny dla danego ośrodka (zwany czasem *dynamicznym współczynnikiem lepkości* w celu odróżnienia od tzw. *kinematycznego współczynnika lepkości*, równego stosunkowi η/ϱ, gdzie ϱ oznacza gęstość ośrodka), Δv — różnicę prędkości obu warstw, czyli prędkość względną jednej warstwy względem drugiej.

Zgodnie z tym, co powiedzieliśmy wyżej, siła lepkości F_l, działająca na każdej z warstw, jest równa co do wartości sile F, lecz przeciwnie skierowana:

$$\mathbf{F} = -\mathbf{F}_l; \qquad F_l = \eta S \frac{\Delta v}{\Delta h}. \tag{9.17}$$

Obie omawiane siły są proporcjonalne do powierzchni warstw cieczy i do tzw. *spadu prędkości* wyrażonego stosunkiem $\Delta v/\Delta h$*.

Kierunek siły F_l na każdej warstwie cieczy jest przeciwny do kierunku prędkości względnej danej warstwy (wektory kreskowane na rys. 9.36). Siła ta dąży do zmniejszenia prędkości warstwy poruszającej się szybciej i do zwiększenia prędkości warstwy poruszającej się wolniej. Dzięki występowaniu tych sił istnieje w przepływającej cieczy dążność do wyrównywania prędkości w różnych obszarach przepływu.

Do głębszego wyjaśnienia mechanizmu zjawiska lepkości powrócimy jeszcze w rozdz. 16, poświęconym fizyce cząsteczkowej, gdzie będziemy mogli podkreślić cechy wspólne zjawisk lepkości, dyfuzji i przewodnictwa cieplnego, obejmowanych wspólną nazwą *zjawisk transportu* (*zjawisk przenoszenia*). Tutaj tylko zasygnalizujemy, że w zjawiskach lepkości, dzięki ruchom cieplnym cząstek cieczy lub gazu, mamy do czynienia z transportem pędu od elementów cieczy poruszających się szybciej do elementów poruszających się wolniej. Właśnie ten transport pędu sprzyja wspomnianemu wyżej wyrównywaniu się prędkości w całym strumieniu przepływającej cieczy lub gazu.

Pytanie. Jakie wielkości fizyczne podlegają transportowi w zjawiskach dyfuzji i przewodnictwa cieplnego? Jakie inne wielkości ulegają przy tym wyrównaniu?

Współczynnik lepkości i jego pomiar. Powróćmy do równania (9.17) i poświęćmy nieco uwagi współczynnikowi lepkości.

Sens fizyczny współczynnika lepkości znajdujemy podstawiając do wzoru (9.17) jednostkowe wartości S, Δv i Δh. W tych warunkach zachodzi odpowiedniość liczbowa współczynnika lepkości η danego ośrodka i siły lepkości F_l. A zatem *współczynnik lepkości danego ośrodka* (w układzie SI) liczbowo wyraża siłę lepkości (w niutonach), powstającą przy ruchu względnym dwóch warstw o jednostkowej powierzchni (1 m^2), jeśli spad prędkości między warstwami jest jednostkowy $\left(1\,\frac{\text{m/s}}{\text{m}} = 1\ \text{s}^{-1}\right)$. *Jednostką*

---

* Tok rozważań dotyczących lepkości cieczy przerwiemy dla wtrącenia uwagi o dość podstawowym znaczeniu. W przypadku granicznym, gdy Δh, mierzone w kierunku największego wzrostu prędkości, zmierza do zera, stosunek $\Delta v/\Delta h$ można zastąpić pochodną dv/dh. Ta pochodna stanowi wartość liczbową wektora zwanego *gradientem prędkości*.

Warto podkreślić, że o gradiencie można mówić w odniesieniu do każdej wielkości φ, która charakteryzuje pewien stan fizyczny istniejący w danym punkcie i zależy od położenia tego punktu w przestrzeni. Gradientem φ w danym punkcie nazywamy wektor zorientowany w kierunku najszybszego wzrastania wielkości φ i mający wartość liczbową równą pochodnej φ względem odległości n mierzonej w kierunku najszybszego wzrostu φ, czyli

$$\mathbf{grad}\,\varphi = \mathbf{l}_n \frac{d\varphi}{dn},$$

gdzie $\mathbf{l}_n$ oznacza wektor jednostkowy w kierunku n. Można więc mówić o gradiencie temperatury przy nierównomiernym ogrzaniu ciała, o gradiencie gęstości przy nierównomiernym rozkładzie masy, o gradiencie potencjału w polu grawitacyjnym lub elektrycznym. Niejednokrotnie będziemy mieli okazję do stosowania tej wielkości.

współczynnika lepkości w układzie SI jest $N \cdot s \cdot m^{-2}$ lub $kg \cdot m^{-1} \cdot s^{-1}$*. Współczynniki lepkości różnych płynów mają bardzo różne wartości, jak to widać z tab. 9.3.

Pomiary współczynników przeprowadza się rozmaitymi metodami. Omówimy przykładowo dwie metody pomiaru współczynnika lepkości, a mianowicie metodę Poiseuille'a opartą na badaniu przepływu cieczy przez rurki włoskowate i metodę Stokesa, opartą na badaniu ruchu ciała kulistego w ośrodku ciekłym lub gazowym.

Tabela 9.3

Dynamiczne współczynniki lepkości płynów

Płyn	Temperatura °C	Dynamiczny współczynnik lepkości 10^{-6} $kg \cdot m^{-1} \cdot s^{-1}$
alkohol etylowy	0	1770
	20	1200
	30	1000
	60	590
eter etylowy	0	300
	20	230
	30	210
rtęć	0	1680
	20	1550
	100	1240
woda	0	1790
	20	1000
	40	650
wodór	0	8,35
	15	8,89
	99,2	10,59
powietrze	0	17,08
	20	17,80
	40	19,04

Poiseuille stwierdził, że średnia prędkość v wypływu cieczy z rurki włoskowatej wyraża się wzorem

$$v = \frac{pR^2}{8\eta l}, \qquad (9.18)$$

gdzie p oznacza ciśnienie, pod którym odbywa się wypływ, R — promień rurki włoskowatej, l — długość rurki. Wzór ten jest słuszny przy niewielkich prędkościach wypływu

* Przed wprowadzeniem układu SI współczynnik lepkości zwykle wyrażano w puazach (P). *Puaz* jest jednostką należącą do układu CGS

$$1\,P = 1\ dyna \cdot s \cdot cm^{-2} = 1\ g \cdot cm^{-1} \cdot s^{-1}.$$

Pytanie. Jaka jest słowna definicja puaza? Jaka jest zależność liczbowa między puazem i jednostką lepkości w układzie SI?

cieczy. Na podstawie tego wzoru można obliczyć objętość V cieczy wypływającej z rurki w czasie t. Uwzględniając kołowy przekrój rurki otrzymujemy

$$V = \pi R^2 v t,$$

czyli — biorąc pod uwagę (9.18) —

$$V = \frac{\pi R^4 p t}{8\eta l},$$

skąd

$$\eta = \frac{\pi R^4 p t}{8 l V}.$$

Wielkości p, R, t, V i l łatwo można wyznaczyć doświadczalnie. Opisana metoda daje możność szybkiego porównywania współczynników lepkości różnych cieczy. Przepuszczamy dwie różne ciecze w jednakowym czasie przez tę samą rurkę włoskowatą pod jednakowym ciśnieniem i mierzymy objętości V_1 i V_2. Spełniają one zależność

$$\frac{\eta_1}{\eta_2} = \frac{V_2}{V_1}.$$

Przechodząc do omówienia drugiej metody wyznaczania współczynnika lepkości, mającej zastosowanie zarówno do cieczy, jak i do gazów, musimy podać pewne dane dotyczące ruchu ciał w ośrodkach lepkich, aczkolwiek temu zagadnieniu jest poświęcony § 9.9. A mianowicie, z badań eksperymentalnych Stokesa wynika, że ciało o kształcie kulistym, spadające w ośrodku lepkim, podlega hamującemu działaniu siły F, skierowanej pionowo w górę i równej

$$F = 6\pi\eta r v, \tag{9.19}$$

gdzie η oznacza współczynnik lepkości danego ośrodka, r — promień spadającej kulki, v — prędkość kulki. Siłę tę będziemy nazywali *siłą Stokesa.* Jest ona proporcjonalna do promienia kulki i do jej prędkości oraz do współczynnika lepkości ośrodka, w którym odbywa się spadanie.

Rys. 9.37

Omówimy dokładniej tę metodę obliczania współczynnika lepkości w odniesieniu do cieczy. Do wysokiego, dość szerokiego naczynia (rys. 9.37), zawierającego badaną ciecz, wrzuca się kulkę o promieniu r i o gęstości ϱ tak dobranej do gęstości cieczy ϱ_1, by spadanie nie odbywało się zbyt szybko.

Na spadającą kulkę działają trzy siły:

1) siła ciężkości

$$P = \tfrac{4}{3}\pi r^3 \varrho g, \tag{9.20}$$

2) siła wyporu wynikająca z prawa Archimedesa

$$W = \tfrac{4}{3}\pi r^3 \varrho_1 g, \tag{9.21}$$

3) siła Stokesa

$$F = 6\pi\eta r v. \tag{9.22}$$

W pierwszym stadium spadania kulki w cieczy prędkość jej rośnie. Równocześnie,

zgodnie ze wzorem (9.19), rośnie siła Stokesa. Przy pewnej wartości prędkości v następuje zrównoważenie się sił,

$$P = W+F, \tag{9.23}$$

i od tej chwili kulka odbywa już ruch jednostajny. Prędkość v tego ruchu znajdujemy doświadczalnie mierząc pewien odcinek drogi s i odpowiadający mu czas przelotu kulki t:

$$v = \frac{s}{t}. \tag{9.24}$$

Po uwzględnieniu wzorów (9.20)–(9.22) oraz (9.24) wyrażenie (9.23) przyjmuje postać

$$\frac{4}{3}\pi r^3 \varrho g = \frac{4}{3}\pi r^3 \varrho_1 g + 6\pi\eta r \frac{s}{t},$$

skąd

$$\eta = \frac{2r^2 g(\varrho-\varrho_1)t}{9s}.$$

Analizując tab. 9.3 można stwierdzić dużą rozpiętość wartości liczbowych współczynników lepkości różnych płynów, jak również zależność współczynnika lepkości określonego płynu od temperatury. Współczynniki lepkości cieczy na ogół maleją ze wzrostem temperatury: w niektórych cieczach, m. in. w wodzie, ten spadek jest nieznaczny, w innych — przede wszystkim w cieczach organicznych — zmiana jest gwałtowna. W przypadku gazów obserwuje się lekki wzrost współczynnika lepkości ze wzrostem temperatury.

9.9. Ruch burzliwy płynów

Badania doświadczalne dotyczące przepływu płynów rzeczywistych przez przewody wykazały, że po przekroczeniu pewnej prędkości granicznej, zależnej od rodzaju płynu (jego gęstości i współczynnika lepkości) i rodzaju przewodu przepływ z laminarnego przekształca się w burzliwy (turbulentny). Ten nowy rodzaj ruchu płynu charakteryzuje to, że cząstki płynu nie poruszają się w kierunku równoległym do osi przewodu, lecz wykonują ruchy nieuporządkowane o różnych kierunkach prędkości. Takiemu zachowaniu się cząstek towarzyszy powstawanie w płynie nieregularnych linii prądu i powstawanie wirów.

Przejście od jednego rodzaju ruchu do drugiego można pokazać w prostym doświadczeniu. Jeśli do strumienia np. wody przepływającej przez szklaną rurę wprowadzić wąską strugę wody zabarwionej, to przy małych prędkościach przepływu wprowadzona struga zachowa swój „nitkowaty" kształt. Przy dużych prędkościach — dzięki nieregularnym prądom i wirom — nastąpi gwałtowne wymieszanie się cieczy zabarwionej i niezabarwionej, w wyniku czego jednolicie zabarwiona ciecz będzie przepływała przez rurę.

Kryterium przejścia ruchu laminarnego w burzliwy zostało ustalone przez Reynoldsa przez wprowadzenie pewnej liczby bezwymiarowej, zwanej obecnie liczbą

Reynoldsa, związanej z gęstością ϱ i z współczynnikiem lepkości η płynu, z jego prędkością oraz z wymiarami przewodu (np. średnicą d przewodu kołowego). Liczba Reynoldsa Re w przypadku przewodu kołowego równa się

$$\mathrm{Re} = \frac{\varrho d v}{\eta}.$$

Przy przepływie płynu przez gładkie proste przewody o przekroju kołowym ruch jest laminarny, gdy $\mathrm{Re} < 2300$, a burzliwy, gdy $\mathrm{Re} > 3000$ (podane wartości są przybliżone). Dla wartości pośrednich przepływ ma charakter niestacjonarny; przy wystąpieniu jakichkolwiek zakłóceń łatwo staje się burzliwy.

Opierając się na definicji liczby Reynoldsa można wyznaczyć *prędkość krytyczną przepływu płynu*, po której przekroczeniu ruch może stać się ruchem burzliwym. Tej krytycznej prędkości v_k odpowiada krytyczna wartość liczby Reynoldsa Re_k, równa 2300. A zatem przy przepływie przez przewód kołowy

$$v_k = \frac{\mathrm{Re}_k \eta}{\varrho d}.$$

Podstawiając wartości liczbowe dla wody znajdziemy

$$v_k = \frac{2{,}3 \cdot 10^{-3}}{d} \frac{\mathrm{m}}{\mathrm{s}},$$

gdy średnica przewodu d jest mierzona w metrach. A zatem w przewodzie kołowym o średnicy np. 2 cm ruch jest laminarny, gdy prędkość przepływu wody nie przekracza 0,115 m/s.

Z ostatniego wzoru widać, że przy pozostałych warunkach niezmienionych prędkość krytyczna jest odwrotnie proporcjonalna do średnicy przewodu.

9.10. Opór ośrodka

Badanie bardzo skomplikowanych zjawisk, towarzyszących ruchowi ciał stałych w ośrodkach lepkich, ma podstawowe znaczenie dla współczesnej komunikacji lądowej, wodnej i powietrznej. Badania takie przeprowadza się zwykle w ten sposób, że zamiast ruchu ciał względem nieruchomego ośrodka wywołuje się ruch ośrodka względem nieruchomego ciała. Takie ujęcie zagadnienia jest możliwe dzięki temu, że na przebieg zjawisk ma wpływ tylko ruch względny wymienionych ciał.

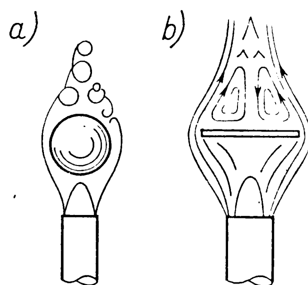
Na plan pierwszy w badaniach wspomnianego typu wysuwa się zagadnienie czynników decydujących o oporze ośrodka. W celu ustalenia wpływu parametrów charakteryzujących ośrodek lepki na opór stosowane są ośrodki o różnych gęstościach, współczynnikach lepkości i prędkościach. Dla ustalenia wpływu parametrów charakteryzujących ciało poruszające się stosuje się ciała modelowe o najrozmaitszych kształtach. Podkreślimy dalej kilka podstawowych wyników badań.

Jeden ze wzorów określających opór ośrodka lepkiego przy ruchu ciała stałego już poznaliśmy w § 9.8. Był to wzór Stokesa na siłę oporu lepkiego, jakiej podlegają ciała kuliste podczas ruchu w ośrodku lepkim. Ze wzoru (9.19) widać, że siła ta jest proporcjonalna do współczynnika lepkości ośrodka, promienia kuli i jej prędkości.

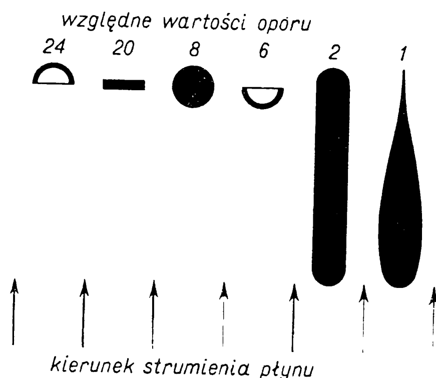
Ten charakter zależności od prędkości (prosta proporcjonalność) występuje tylko przy prędkościach małych. Przy prędkościach dużych opór jest w przybliżeniu proporcjonalny do drugiej potęgi prędkości. Przy prędkościach większych od prędkości głosu w powietrzu zmieniają się radykalnie warunki ruchu w ośrodkach gazowych, gdyż już wtedy nie można zaniedbywać ściśliwości gazu.



Rys. 9.38



Rys. 9.39



Rys. 9.40

Opór ośrodka jest w dużym stopniu zależny od kształtu poruszającego się ciała, w szczególności od kształtu przedniej i tylnej powierzchni. Bardzo niekorzystnym zjawiskiem towarzyszącym ruchowi ciał w ośrodku lepkim jest tworzenie się wirów. Przykład tworzenia się wirów za poruszającą się płytą widoczny jest na rys. 9.38. Tworzenie się wirów w otoczeniu ciała stałego umieszczonego w przepływającym gazie można obserwować w prostym doświadczeniu. Do płomienia palnika gazowego wstawiamy ciała o różnych kształtach, np. kule, płytki itp. Na cieniu obserwuje się wiry powstające za tymi ciałami (schematycznie podane na rys. 9.39a i b).

Tworzeniu się wirów towarzyszy wzrost oporu ośrodka i bezużyteczne zużywanie się energii, która następnie rozprasza się w ośrodku w postaci ciepła. Zmniejszenie liczby wirów — mające podstawowe znaczenie w zagadnieniach komunikacji — osiąga się przez odpowiedni dobór kształtu poruszającego się ciała. Kształt tzw. opływowy po-

dobny do kształtu spadającej kropli stwarza najkorzystniejsze warunki ruchu. Taki kształt często występuje w przyrodzie (ciało ryb, ptaków itp.), jak również często nadawany jest pojazdom przeznaczonym do ruchu na lądzie, morzu i w powietrzu (łodzie, statki, samochody, samoloty, rakiety itp.). Rys. 9.40 ilustruje względne zmiany oporu ośrodka podczas ruchu ciał o jednakowych rozmiarach poprzecznych, lecz różnym kształcie lub różnym ustawieniu, czy też różnym kierunku ruchu. Wszystkie wartości odpowiadają tej samej prędkości ruchu.

Jako przykład wpływu powietrza na ruch ciał spadających z dużych wysokości rozpatrzmy opadanie spadochroniarza. Jeśli z własnej woli skoczka lub z racji awarii spadochron się nie otwiera, to jednak wskutek oporu ośrodka ruch jest przyspieszony tylko przez 10–12 pierwszych sekund ruchu. W tym czasie prędkość gwałtownie rośnie, a wraz z nią rośnie opór. W końcu dochodzi do zrównoważenia siły ciężkości i siły oporu i odtąd już spadochroniarz odbywa ruch jednostajny z prędkością około 60 m/s. (Tego samego rzędu są prędkości spadających na Ziemię meteorów.) Otwarcie spadochronu od razu na początku spadania zwiększa opór, a tym samym zmniejsza prędkość ruchu jednostajnego. Rozmiary spadochronów są tak dobierane, aby prędkość końcowa ruchu, czyli prędkość ruchu jednostajnego wynosiła około 2 m/s.

Pytania i zadania

1. Podać ogólną charakterystykę sprężystych właściwości ciał ciekłych i gazowych. Jakie są cechy wspólne, a jakie różnice?

2. Dlaczego w prawie Pascala ważne są założenia, że ciecz jest nieważka i nieściśliwa?

3. Jaka jest różnica między parciem i ciśnieniem? W jakich jednostkach w układzie SI wyrażamy te wielkości?

4. Jaki jest warunek równowagi cieczy w naczyniach połączonych?

5. Jaka jest treść prawa Archimedesa w odniesieniu do płynów?

6. Jakie są warunki pływania ciał?

7. Wykazać, że ciężar właściwy γ_s ciała stałego wyraża się wzorem

$$\gamma_s = \frac{P_1}{P_1 - P_2}\gamma,$$

gdzie P_1 oznacza ciężar badanego ciała w powietrzu, P_2 — ciężar odważników równoważących to ciało po zanurzeniu go w cieczy o znanym ciężarze właściwym γ.

8. Wykazać, że ciężar właściwy γ_c cieczy wyraża się wzorem

$$\gamma_c = \frac{P_1 - P_3}{P_1 - P_2}\gamma,$$

gdzie P_1 oznacza ciężar dowolnego pomocniczego ciała stałego w powietrzu, P_2 — ciężar odważników równoważących to ciało po zanurzeniu go w cieczy o znanym ciężarze właściwym γ, P_3 — ciężar odważników równoważących to ciało po zanurzeniu go w cieczy o ciężarze właściwym badanym.

9. Omówić zasadę budowy i działania areometrów.

10. Podać treść prawa Boyle'a–Mariotte'a. Ściśle sformułować warunki jego stosowalności.

11. Czym się zajmuje hydro- i aerodynamika?

12. Scharakteryzować przepływ a) stacjonarny, b) laminarny, c) burzliwy.

13. Omówić dokładnie a) strumień masy i objętości, b) gęstość strumienia masy i objętości, c) związek gęstości strumienia objętości z prędkością przepływu przez dany przekrój i z gęstością linii prądu.

14. Podać definicje linii prądu i rurek prądu. Na czym polega praktyczne znaczenie modelu rurek prądu?

15. Jaki jest sens fizyczny równań ciągłości strumienia masy i objętości?

16. Jakie założenia obowiązują przy prawie Bernoulliego? Jaki jest tok rozumowania przy jego wyprowadzaniu?

17. Dzięki czemu można stosować prawo Bernoulliego do gazów przy prędkościach przepływu mniejszych od prędkości głosu? Podać przykłady zastosowań tego prawa.

18. Czy prawo Bernoulliego z równą dokładnością stosuje się do przepływu wody, gliceryny i powietrza przez przewody?

19. Jakie fakty świadczą o istnieniu lepkości w cieczach i gazach? Od czego zależy siła lepkości przy ruchu laminarnym cieczy? Jaki jest jej kierunek?

20. Jaki jest sens fizyczny współczynnika lepkości? Określić jednostkę współczynnika lepkości w układzie SI? Czy ma ona własną nazwę?

21. Na jakich prawach opierają się opisane w tym rozdziale metody pomiaru współczynnika lepkości? Która z metod pozwala na szybkie wyznaczanie względnego współczynnika lepkości cieczy? Która z metod może być stosowana do gazów?

22. Jaka jest zależność oporu ośrodka: a) od prędkości ruchu ciał, b) od kształtu ciał?

23. Co można przyjmować za kryterium przejścia ruchu laminarnego w burzliwy?

24. Dlaczego we współczesnej komunikacji tak dużo uwagi poświęca się problemowi zmniejszenia możliwości tworzenia się wirów?

25. Jednorodna kula o objętości V i o ciężarze właściwym γ pływa na granicy dwóch nie mieszających się cieczy. Ciężar właściwy górnej cieczy jest γ_1, dolnej γ_2. Jaka część objętości kuli zanurzona jest w górnej, a jaka w dolnej cieczy?

Odp. $\frac{\gamma_2-\gamma}{\gamma_2-\gamma_1}$, $\frac{\gamma-\gamma_1}{\gamma_2-\gamma_1}$.

26. Areometr o masie $m = 63$ g zanurza się w wodzie do podziałki oznaczonej cyfrą 1. W jakiej odległości od tej podziałki znajdują się kreski odpowiadające głębokościom zanurzeń areometru w cieczach o gęstościach 0,9 g/cm^3 oraz 1,1 g/cm^3? Powierzchnia przekroju rurki, na której znajduje się podziałka, $S = 2{,}1$ cm^2.

Odp. 3,3 cm, 2,7 cm.

27. W dnie naczynia znajduje się otwór, przez który wypływa woda. Wysokość słupa wody w naczyniu równa się h (zakładamy, że jest stała). Z jaką prędkością wypływa woda, jeśli 1) naczynie jest nieruchome, 2) porusza się ruchem: a) jednostajnym, b) jednostajnie przyspieszonym do góry z przyspieszeniem a, c) jednostajnie przyspieszonym w dół z przyspieszeniem a.

Odp. 1) $\sqrt{2gh}$, 2b) $\sqrt{2h(g+a)}$, 2c) $\sqrt{2h(g-a)}$.

28. W szerokiej części poziomo ułożonego rurociągu woda płynie z prędkością 50 cm/s pod ciśnieniem 1,5 at. W wąskiej części przewodu ciśnienie wynosi 1,4 at. Jaka jest prędkość przepływu wody w wąskiej części przewodu?

Odp. 4,5 m/s.

29. Powierzchnia tłoka strzykawki wynosi $S_1 = 1{,}2$ cm^2, a powierzchnia przekroju igły $S = 1$ mm^2. Ile czasu trwać będzie wypływ wody ze strzykawki, jeśli na tłok działa siła 0,5 kG na odcinku długości $l = 4$ cm.

Odp. ok. 0,53 s.

ROZDZIAŁ 10

Drgania mechaniczne

Pojęcie drgań jest szeroko w fizyce stosowane. Przykładowo można wymienić drgania mechaniczne, akustyczne, elektromagnetyczne. Aczkolwiek w rozdziale niniejszym ograniczymy się do omawiania *drgań mechanicznych*, to jednak, o czym się przekonamy z rozdziałów dalszych, wiele pojęć i wniosków można będzie uogólnić na drgania niemechaniczne.

Analizując mechaniczne drgania ciała dookoła położenia równowagi, uwagę naszą skoncentrujemy z konieczności na kilku wybranych zagadnieniach. Przede wszystkim zajmiemy się *drganiami mechanicznymi harmonicznymi*, zwanymi także ruchem harmonicznym (równanie tego ruchu podaliśmy w § 1.3). Następnie zajmiemy się charakterystyką *drgań* tzw. *tłumionych* na przykładzie drgań powstających wtedy, gdy ciało wykonujące ruch harmoniczny jest poddane działaniu dodatkowej siły hamującej ruch, proporcjonalnej do prędkości chwilowej i skierowanej do niej przeciwnie. Wiele sił hamujących, występujących w przyrodzie (np. siła lepkości), ma właśnie taki charakter. Dalej omówimy *drgania* tzw. *wymuszone* na przykładzie drgań powstających wtedy, gdy na ciało wykonujące ruch harmoniczny działa dodatkowa siła sinusoidalnie zmienna w czasie. Przy tej okazji przeanalizujemy bardzo ważne i rozpowszechnione zjawisko rezonansu. Na zakończenie wspomnimy o *drganiach* tzw. *samowzbudnych*, nieco więcej uwagi poświęcając *drganiom samowzbudnym relaksacyjnym.*

10.1. Ruch harmoniczny

Z punktu widzenia matematyki ruch harmoniczny punktu materialnego opisuje równanie

$$s = A\sin\omega t, \tag{10.1}$$

gdzie s oznacza wychylenie punktu drgającego od położenia równowagi, t — czas, A i ω — wielkości stałe w danym ruchu, tzn. niezależne od czasu.

Następujące uwagi pozwolą zrozumieć sens fizyczny obu wymienionych stałych. Znaczenie stałej A wynika z charakteru funkcji sinus: funkcja ta może się zmieniać

w granicach od -1 do $+1$, a zatem wychylenie s od położenia równowagi może się zmieniać w granicach

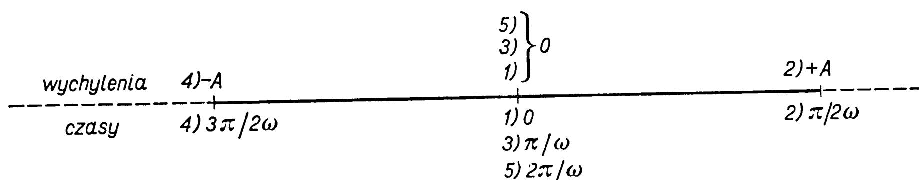
$$-A \leqslant s \leqslant +A.$$

Innymi słowy, punkt drgający może się odsuwać od położenia równowagi najdalej o $\pm A$. Stała A oznacza więc największe wychylenie od położenia równowagi, zwane *amplitudą ruchu harmonicznego.*

Zbadajmy zmiany wychylenia od położenia równowagi w zależności od czasu. W tabelce podanej niżej kolejnym rozważanym chwilom przypisujemy takie wartości czasu t, aby argument sinusa we wzorze (10.1) przyjmował wartości 0, $\pi/2$, π, $3\pi/2$, 2π.

Kolejna chwila	t	ωt	s
1	0	0	0
2	$\pi/2\omega$	$\pi/2$	$+A$
3	$2\pi/2\omega$	π	0
4	$3\pi/2\omega$	$3\pi/2$	$-A$
5	$4\pi/2\omega$	2π	0

Przedstawimy tę zależność na osi liczbowej, zaznaczając nad osią wartości wychyleń od położenia równowagi (punktu 0), poniżej zaś osi — wartości czasu (rys. 10.1). W chwili $t = 0$ punkt drgający przechodzi przez położenie równowagi, następnie oddala się od niego w prawo osiągając w chwili $t = \pi/2\omega$ wychylenie maksymalne równe $+A$, po czym stopniowo wraca do położenia równowagi, mija go w chwili $t = \pi/\omega$ i przesuwa się dalej na lewo osiągając ponownie w chwili $t = 3\pi/2\omega$ maksymalne



Rys. 10.1

wychylenie równe $-A$, skąd w chwili $t = 2\pi/\omega$ wraca do położenia równowagi. Począwszy od tego położenia powtarzają się wszystkie poprzednie fazy ruchu w tym samym rytmie. Zasadniczą cechą ruchu harmonicznego jest więc jego okresowość. Jak widać z tabelki i rys. 10.1, czas trwania jednego pełnego drgnienia T, zwany *okresem*, wynosi

$$T = \frac{2\pi}{\omega}. \tag{10.2}$$

Częstotliwość ruchu ν, czyli liczba pełnych drgań dokoła położenia równowagi, wykonywanych w jednostce czasu, jest odwrotnością okresu:

$$\nu = \frac{1}{T} = \omega/2\pi. \tag{10.3}$$

Ze wzorów (10.2) i (10.3) wynika znaczenie stałej ω, zwanej *pulsacją* (częstością kątową):

$$\omega = \frac{2\pi}{T} \quad \text{lub} \quad \omega = 2\pi\nu. \tag{10.4}$$

W układzie SI *jednostką pulsacji* jest radian na sekundę (rad/s).

Uwzględniając (10.4) można równanie ruchu harmonicznego napisać w postaci

$$s = A\sin\frac{2\pi}{T}t. \tag{10.5}$$

Jak można łatwo sprawdzić, na pozór dość sztucznie dobrane czasy uwzględnione w tabelce zależności s od t odpowiadają kolejno wartościom 0, $T/4$, $2T/4$, $3T/4$, T.

Drgania dokoła położenia równowagi odbywające się zgodnie z równaniami (10.1) lub (10.5) nazywamy często *oscylacjami harmonicznymi*, a ciało wykonujące takie drgania — *oscylatorem harmonicznym*.

Z kolei przejdźmy do obliczenia prędkości v ruchu harmonicznego. Pamiętając, że prędkość jest pochodną drogi względem czasu, znajdujemy

$$v = \frac{ds}{dt} = A\omega\cos\omega t = A\omega\cos\frac{2\pi}{T}t. \tag{10.6}$$

Jak widać, ***prędkość jest wielkością zmienną, okresową***. Ponieważ $-1 \leqslant \cos\omega t \leqslant +1$, więc największa wartość prędkości $v_{max} = A\omega$.

Zbadajmy zmienność prędkości w zależności od czasu uwzględniając te same kolejne chwile, co i poprzednio:

Kolejna chwila	t	v	a
1	0	$A\omega$	0
2	$T/4$	0	$-A\omega^2$
3	$2T/4$	$-A\omega$	0
4	$3T/4$	0	$+A\omega^2$
5	T	$A\omega$	0

Rysunek 10.2 przedstawia zestawienie tych prędkości na osi liczbowej, uwidoczniające równocześnie ich powiązanie z wychyleniami od położenia równowagi. W punktach odpowiadających największym wychyleniom ciało ma prędkość równą zeru (tzn. na chwilę się zatrzymuje). W chwili mijania położenia równowagi prędkość jest największa co do wartości bezwzględnej, równa $+A\omega$ lub $-A\omega$, tzn. skierowana w prawo lub w lewo od położenia równowagi.

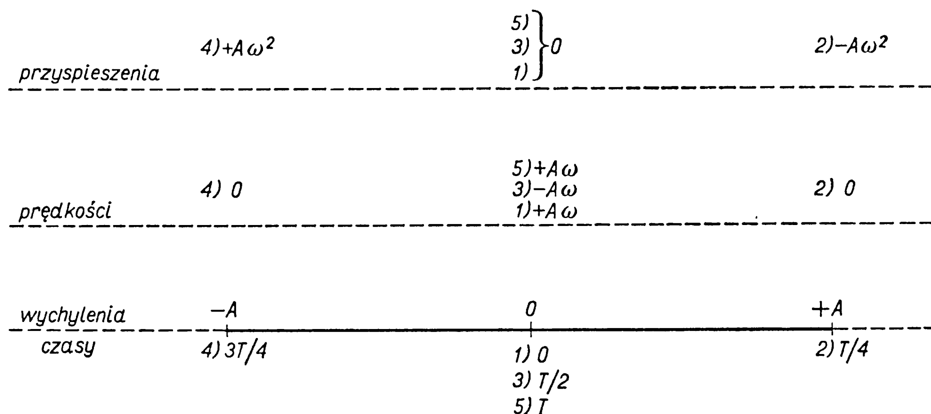
W analogiczny sposób badamy przyspieszenie. Pamiętając, że przyspieszenie a jest pochodną prędkości względem czasu, znajdujemy

$$a = \frac{dv}{dt} = -A\omega^2\sin\omega t. \tag{10.7}$$

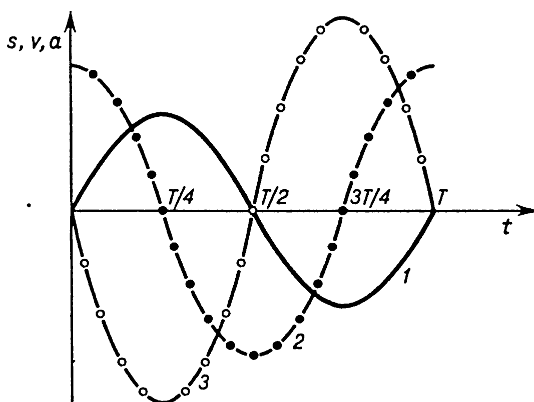
Ale iloczyn $A \sin \omega t$ wyraża wychylenie s od położenia równowagi, a zatem

$$a = -\omega^2 s \quad \text{lub} \quad a = -\frac{4\pi^2}{T^2} s. \tag{10.8}$$

Innymi słowy, ***przyspieszenie w ruchu harmonicznym jest proporcjonalne do wychylenia od położenia równowagi.*** Jest to wniosek wypływający z przyjętej przez nas definicji ruchu harmonicznego, zawartej w równaniu (10.1). Jest to jednak równocześnie tak charakterystyczna cecha ruchu harmonicznego, że często na podstawie występowania tej cechy zalicza się badany ruch do ruchów harmonicznych (por. § 10.2).



Rys. 10.2



Rys. 10.3

Powróćmy jeszcze do wzoru (10.8). Znak minus w tym wzorze przypomina, że ***zwrot przyspieszenia jest zawsze przeciwny do zwrotu wychylenia od położenia równowagi***: gdy ciało znajduje się po prawej stronie licząc od położenia równowagi, przyspieszenie jest skierowane w lewo i odwrotnie. Uwzględniając równocześnie zwroty prędkości można łatwo określić, na jakich odcinkach drogi ruch jest przyspieszony lub opóźniony.

Zbadajmy zależność przyspieszenia od czasu wprowadzając do ostatniej tabelki końcową kolumnę, zawierającą wartości a, wynikające ze wzoru (10.8). Otrzymane wartości przyspieszenia uwzględniamy też na rys. (10.2) uzyskując w ten sposób obraz zmian przyspieszenia w zależności od czasu, wychylenia od położenia równowagi i prędkości. Tak np. z rys. 10.2 widać, że punkt drgający ma największe przyspieszenie wtedy, gdy jego wychylenie od położenia równowagi jest maksymalne, a prędkość równa zeru. Wyniki naszych rozważań dotyczących kinematyki ruchu harmoniczego można też przedstawić na rys. 10.3. Krzywe *1*, *2*, *3* przedstawiają odpowiednio zależność s, v i a od czasu w ruchu opisanym równaniem (10.1).

Zbadaliśmy ruch harmoniczny wyrażający się równaniem:

$$s = A \sin \omega t.$$

Można wykazać, że równania typu

$$s = A \sin(\omega t + \varphi), \tag{10.9a}$$

$$s = A \cos(\omega t + \varphi) \tag{10.9b}$$

również przedstawiają ruch harmoniczny. W przypadku odpowiadającym zależności (10.9a) w chwili $t = 0$, tzn. w chwili rozpoczęcia rachuby czasu, ciało nie znajduje się w położeniu równowagi, lecz ma już wychylenie $s_0 = A \sin \varphi$. Wartość kąta φ nazywamy *fazą początkową* ruchu harmonicznego, podczas gdy całość wyrażenia $(\omega t + \varphi)$ nazywamy *fazą ruchu harmonicznego* albo *fazą drgania*. A zatem ruchy harmoniczne opisane równaniami (10.1) i (10.9a) różnią się tylko fazą początkową.

P y t a n i e. Czy można różnicę ruchów harmonicznych (10.9a) i (10.9b) sprowadzić także do różnicy faz początkowych?

Dotychczasowe rozważania dotyczyły kinematyki ruchu harmonicznego. Obecnie przejdziemy do ujęcia tego ruchu z punktu widzenia dynamiki.

Z istnienia przyspieszenia wynika, że punkt materialny o masie m, wykonujący ruch harmoniczny, podlega działaniu siły

$$F = -m\omega^2 s = -m \frac{4\pi^2}{T^2} s.$$

Siła ta jest proporcjonalna do wychylenia i przeciwnie do niego skierowana (taki charakter mają m. in. siły sprężyste). Współczynnik proporcjonalności $m\omega^2 = k$ nazywany jest zwykle współczynnikiem sprężystości, aczkolwiek ze względu na to, że drgania harmoniczne mogą być wywołane nie tylko przez siły sprężyste, lepiej byłoby nazywać go współczynnikiem quasi-sprężystości.

Po podstawieniu symbolu współczynnika otrzymujemy

$$F = -ks.$$

Uwzględniając drugą zasadę dynamiki można to równanie napisać w postaci

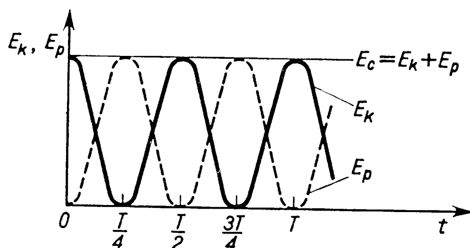
$$m \frac{d^2 s}{dt^2} = -ks \quad \text{lub} \quad m \frac{d^2 s}{dt^2} + ks = 0. \tag{10.10}$$

Całość rozważań dotyczących ruchu harmonicznego można oprzeć na równaniu różniczkowym (10.10) jako równaniu wyjściowym definiującym badany ruch. Rozwiązując to równanie otrzymalibyśmy wyrażenie na s typu (10.9a). W naszym ujęciu zastosowaliśmy postępowanie odwrotne: od równania charakteryzującego zależności kinematyczne przeszliśmy do równania (10.10) charakteryzującego zależności dynamiczne. Warto jednak zapamiętać, że rozwiązaniem równania różniczkowego typu (10.10), które, jak się okaże, opisuje szereg różnorodnych zjawisk fizycznych, jest równanie typu (10.9a).

Przejdźmy teraz do rozpatrzenia ruchu harmonicznego z punktu widzenia energii ciała drgającego. Ciało wykonujące ruch harmoniczny ma energię kinetyczną E_k

$$E_k = \tfrac{1}{2}mA^2\omega^2\cos^2(\omega t+\varphi).$$

Energia kinetyczna jest więc podczas ruchu zmienna. Zmienność energii kinetycznej w ciągu jednego okresu ruchu harmonicznego przedstawiono na rys. 10.4 linią ciągłą (rysunek dotyczy ruchu $s = A\sin\omega t$).



Rys. 10.4

Energia potencjalna E_p ciała wykonującego ruch harmoniczny równa się pracy, którą ciało drgające może wykonać wracając od wychylenia s do położenia równowagi. Wobec zmienności siły pracę W znajdujemy przez całkowanie:

$$E_p = W = \int_s^0 F\,ds = -\int_s^0 ks\,ds = k\int_0^s s\,ds = \frac{ks^2}{2}.$$

Podstawiając wartości s i k otrzymujemy

$$E_p = \tfrac{1}{2}m\omega^2A^2\sin^2(\omega t+\varphi).$$

Jak widać z tego wzoru, energia potencjalna ciała wykonującego ruch harmoniczny również zmienia się w czasie wykonywania ruchu. Zmienność energii potencjalnej w ciągu jednego okresu ruchu harmonicznego (przy $\varphi = 0$) przedstawiono na rys. 10.4 linią kreskowaną.

Gdy ruch odbywa się bez żadnych strat energii na pokonywanie oporów, całkowita energia E_c, czyli suma energii kinetycznej i potencjalnej wyraża się wzorem

$$E_c = E_k+E_p = \tfrac{1}{2}m\omega^2A^2[\cos^2(\omega t+\varphi)+\sin^2(\omega t+\varphi)],$$

czyli

$$E_c = \tfrac{1}{2}m\omega^2A^2. \tag{10.11}$$

Widzimy więc, że w przypadku drgań harmonicznych o określonej pulsacji *całkowita energia mechaniczna jest proporcjonalna do kwadratu amplitudy*. Załóżmy, że tak dobraliśmy początek rachuby czasu, że faza początkowa $\varphi = 0$. Wtedy w chwili $t = 0$ lub $t = T/2$ $E_p = 0$, podczas gdy $E_k = \frac{1}{2}m\omega^2 A^2$, a w chwili $t = T/4$ lub $t = 3T/4$ $E_k = 0$, a $E_p = \frac{1}{2}m\omega^2 A^2$. Innymi słowy, wartości energii kinetycznej i potencjalnej wahają się między zerem a wartością maksymalną, lecz *całkowita energia mechaniczna ciała wykonującego drganie harmoniczne jest stała i równa*

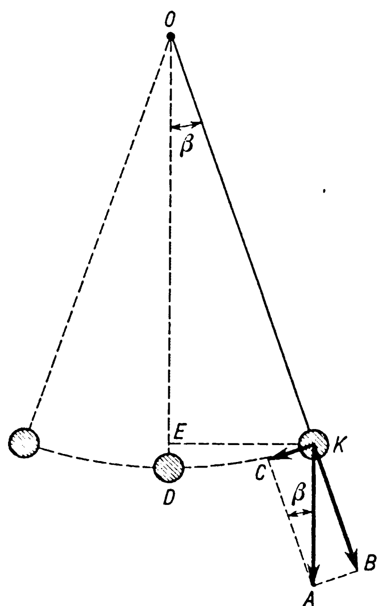
$$E_c = \frac{1}{2}m\omega^2 A^2.$$

Stałość całkowitej energii E_c w ruchu harmonicznym jest też widoczna na rys. 10.4.

10.2. Wahadło matematyczne

Zastanówmy się, czy ruch wahadła matematycznego można traktować jako ruch harmoniczny.

Wahadło matematyczne jest to punkt materialny (np. w postaci kulki o masie m i bardzo małym promieniu) zawieszony na nieważkiej i nierozciągliwej nici (rys. 10.5). Wychylając nić o niewielki kąt β od położenia pionowego i puszczając swobodnie kulkę K wywołujemy jej drgania dokoła położenia równowagi D. Aczkolwiek w praktyce amplituda tych drgań wskutek pokonywania oporów ruchu stopniowo maleje, okres wahań można uważać za stały. Tę własność ruchu wahadła nazywamy izochronizmem. W dalszych rozważaniach pomijać będziemy siły oporu, zakładając, że na kulkę działa tylko siła ciężkości $KA = mg$. Siłę tę rozkładamy na dwie składowe. Jedna z nich, KB, działa wzdłuż nici powodując tylko jej naprężenie, druga, KC, styczna do toru wahadła, wywołuje jego ruch z przyspieszeniem a, a zatem równa się ma. Z prostych zależności geometrycznych wynika, że



Rys. 10.5

$$\sphericalangle EOK = \sphericalangle CAK = \beta.$$

Z trójkąta CAK

$$\sin\beta = ma/mg.$$

Z trójkąta OKE

$$\sin\beta = EK/l,$$

gdzie l oznacza długość wahadła, czyli odległość od punktu zawieszenia O do środka kulki.

A zatem

$$\frac{a}{g} = \frac{EK}{l}, \quad \text{czyli} \quad a = \frac{g}{l}EK. \qquad (10.12)$$

Gdy kąty wychylenia nici od położenia pionowego są małe, nie przekraczają 5–6°, można w przybliżeniu traktować odcinek EK jako równy łukowi DK, czyli równy wychyleniu kulki od położenia równowagi. A zatem — wobec stałości czynnika g/l w okre-

ślonym miejscu na Ziemi — przy małych wychyleniach wahadła zachodzi proporcjonalność przyspieszenia do wychylenia. Poza tym przyspieszenie (a zatem i siła) ma zwrot do położenia równowagi. Ale te cechy wymienialiśmy jako charakterystyczne dla ruchu harmonicznego, możemy zatem ruch wahadła matematycznego w przypadku małych wychyleń od położenia równowagi uważać za ruch harmoniczny.

Z porównania wartości współczynników proporcjonalności we wzorach (10.8) i (10.12) wynika zależność

$$\frac{g}{l} = \frac{4\pi^2}{T^2},$$

z której otrzymuje się znany wzór na okres wahadła matematycznego:

$$T = 2\pi\sqrt{\frac{l}{g}}. \tag{10.13}$$

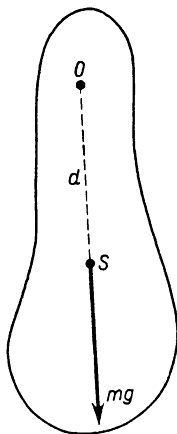
Przekształcając to wyrażenie można też dojść do postaci

$$g = \frac{4\pi^2 l}{T^2},$$

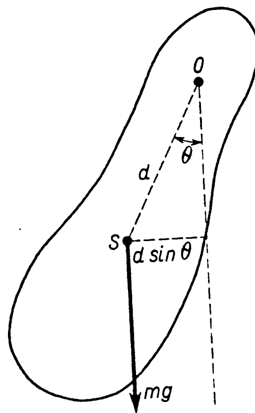
przydatnej przy wyznaczaniu przyspieszenia ziemskiego metodą wahadła matematycznego.

10.3. Wahadło fizyczne

Wahadłem fizycznym nazywamy bryłę sztywną dowolnego kształtu, zawieszoną tak, że może się wahać dookoła pewnej osi przechodzącej przez tę bryłę. Ruch wahadła fizycznego może być wywołany działaniem różnych sił. Prawa rządzące ruchem wahadła fizycznego omówimy na przykładzie wahadła fizycznego grawitacyjnego.



Rys. 10.6



Rys. 10.7

Wahadło fizyczne grawitacyjne przedstawia rys. 10.6. Jest to bryła sztywna dowolnego kształtu o środku ciężkości w punkcie S, zawieszona w ten sposób, że może się obracać bez tarcia dookoła osi poziomej, przechodzącej przez punkt O. Rys. 10.6 przed-

stawia bryłę w położeniu równowagi: środek ciężkości S leży na pionie przechodzącym przez punkt O. Odległość OS od środka ciężkości do osi obrotu oznaczmy przez d, masę bryły przez m, moment bezwładności bryły względem osi obrotu przez I.

Na rysunku 10.7 wahadło jest już wychylone od położenia równowagi. Miarą wychylenia jest kąt θ zaznaczony na rysunku. W tym nowym położeniu na wahadło działa moment siły ciężkości, równy liczbowo $mgd\sin\theta$, skierowujący wahadło w stronę położenia równowagi, co uwzględniamy traktując moment jako wielkość ujemną.

Z teorii ruchu obrotowego wiemy, że moment siły wywołującej ten ruch równa się iloczynowi momentu bezwładności i przyspieszenia kątowego α. A zatem w rozważanym przypadku wahadła fizycznego grawitacyjnego zachodzi równość

$$I\alpha = -mgd\sin\theta. \tag{10.14}$$

Iloczyn mgd jest maksymalną wartością momentu siły ciężkości odpowiadającą wychyleniu wahadła o kąt $\theta = 90°$ od położenia równowagi. Nazywamy ją *momentem kierującym wahadła grawitacyjnego* i oznaczamy literą D:

$$D = mgd. \tag{10.15}$$

Ograniczając rozważania ruchu wahadła fizycznego do bardzo małych kątów wychylenia można skorzystać z przybliżonej równości $\sin\theta \approx \theta$. Wtedy po uwzględnieniu (10.15) wyrażenie (10.14) przyjmuje postać

$$I\alpha = -D\theta$$

lub po uwzględnieniu tego, że z definicji $\alpha = d^2\theta/dt^2$

$$\frac{d^2\theta}{dt^2} = -\frac{D}{I}\theta. \tag{10.16}$$

Równanie to jest analogiczne do równania ruchu harmonicznego (10.8), z tym że zamiast wychylenia liniowego występuje wychylenie kątowe i stosunek D/I zastępuje czynnik ω^2 lub $4\pi^2/T^2$.

$$\frac{D}{I} = \frac{4\pi^2}{T^2},$$

skąd

$$T = 2\pi\sqrt{\frac{I}{D}}. \tag{10.17}$$

Wzór powyższy, aczkolwiek wyprowadzony dla wahadła fizycznego grawitacyjnego, jest słuszny dla dowolnego wahadła fizycznego, z tym że oczywiście w zależności od warunków powstawania ruchu wchodzą w grę momenty kierujące różnych sił (np. sprężystych, magnetycznych itp.).

Wahadło matematyczne można uważać za przypadek szczególny wahadła fizycznego grawitacyjnego. Podstawiając $I = ml^2$ i $D = mgl$ otrzymujemy znany wzór na okres wahadła matematycznego:

$$T = 2\pi\sqrt{\frac{l}{g}}.$$

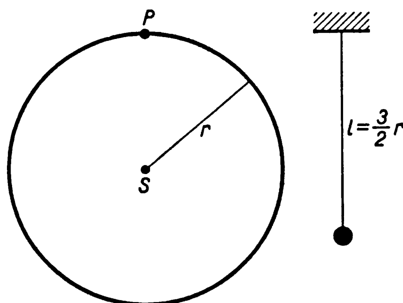
Z wahadłem fizycznym często się kojarzy pojęcie tzw. długości zredukowanej. Przez *długość zredukowaną wahadła fizycznego* rozumiemy długość wahadła matematycznego mającego ten sam okres wahań, a więc spełniającego warunek

$$T_{\text{fiz}} = T_{\text{mat}},$$

czyli $I/D = l_{\text{zred}}/g$, skąd

$$l_{\text{zred}} = \frac{Ig}{D} = \frac{I}{md},$$

gdzie d oznacza odległość od osi obrotu do środka ciężkości wahadła fizycznego.



Rys. 10.8

Tak np. można obliczyć, że dysk o promieniu r, zawieszony w punkcie P znajdującym się na jego obwodzie, ma długość zredukowaną równą $\frac{3}{2}r$. Ilustruje to rys. 10.8.

P y t a n i e. Czy długość zredukowana wahadła fizycznego zależy od wyboru osi obrotu?

10.4. Drgania harmoniczne tłumione

W paragrafie 10.1 rozważaliśmy drganie harmoniczne o stałej amplitudzie i niezmiennej energii całkowitej E_c. Z punktu widzenia dynamiki ruch taki — nazwiemy go ruchem harmonicznym nietłumionym — opisany jest równaniem

$$m\frac{d^2s}{dt^2} = -m\omega_0^2 s \quad \text{lub} \quad m\frac{d^2s}{dt^2} = -ks,$$

gdzie k oznacza współczynnik sprężystości równy iloczynowi $m\omega_0^2$. (Warto zwrócić uwagę, że w stosunku do symboli używanych w § 10.1 zaszła pewna zmiana, a mianowicie symbol ω_0 w tym i następnych paragrafach używany jest zamiast ω stosowanego w § 10.1 i oznacza *pulsację drgań harmonicznych nietłumionych*, zwaną też *pulsacją drgań własnych*.)

Drgania odbywane w warunkach rzeczywistych, w dowolnym ośrodku materialnym, zawsze są połączone z przekazywaniem energii otoczeniu w związku z pokonywaniem sił oporu. W wyniku wykonywanej pracy energia ciała drgającego maleje, zmniejsza się też amplituda drgań. Drgania nie podtrzymywane siłą zewnętrzną ulegają tłumieniu, gasną, zanikają — stąd ich nazwy: *drgania tłumione, gasnące, zanikające*.

Rozpatrzmy przypadek zanikania drgań harmonicznych pod wpływem siły hamującej **F** proporcjonalnej do prędkości i skierowanej do niej przeciwnie:

$$F = -bv = -b\frac{ds}{dt}, \tag{10.18}$$

gdzie b oznacza współczynnik proporcjonalności, zwany współczynnikiem oporu. Wymieniony przypadek jest dość rozpowszechniony, gdyż siły lepkości i oporu ośrodka mają — przy niewielkich prędkościach — właśnie taki charakter.

Uwzględnienie siły (10.18) w równaniu ruchu harmonicznego odbywającego się wzdłuż osi x dokoła położenia równowagi $x = 0$ prowadzi do równania

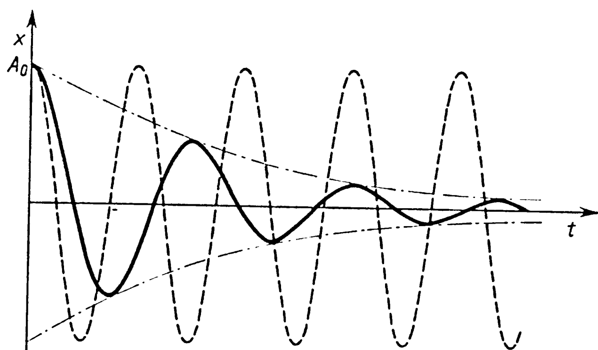
$$m\frac{d^2x}{dt^2} = -kx - b\frac{dx}{dt}.$$

Dzieląc obie strony równania przez masę i grupując wszystkie wyrazy po jednej stronie otrzymujemy

$$\frac{d^2x}{dt^2} + \frac{b}{m}\frac{dx}{dt} + \frac{k}{m}x = 0.$$

Wprowadzając oznaczenie $b/m = 2\delta$ i podstawiając zamiast stosunku k/m jego wartość ω_0^2 otrzymujemy równanie różniczkowe ruchu harmonicznego tłumionego w zwykle mu nadawanej postaci

$$\frac{d^2x}{dt^2} + 2\delta\frac{dx}{dt} + \omega_0^2 x = 0. \tag{10.19}$$



Rys. 10.9

Można sprawdzić przez dwukrotne różniczkowanie (10.20), że w przypadku bardzo małych wartości δ rozwiązaniem równania różniczkowego (10.19) jest wyrażenie

$$x = A_0 e^{-\delta t}\sin(\omega t + \varphi), \tag{10.20}$$

gdzie x oznacza wychylenie od położenia równowagi, A_0 — początkową, maksymalną amplitudę, e — podstawę logarytmów naturalnych, δ — współczynnik zwany stałą tłumienia, ω — pulsację drgań tłumionych (por. (10.21)). Z równania (10.20) widać,

że *amplituda A drgań tłumionych zmienia się wykładniczo z biegiem czasu* przyjmując wartość zerową teoretycznie po czasie nieskończenie długim:

$$A = A_0 e^{-\delta t}.$$

Rysunek 10.9 przedstawia przebieg drgań tłumionych (linia ciągła) w zestawieniu z drganiami nietłumionymi o amplitudzie A_0 (linia kreskowana). Linie utworzone z kropek i kresek są liniami charakteryzującymi wykładniczy spadek amplitudy drgań tłumionych z biegiem czasu.

Wykres drgań zarówno tłumionych, jak i nietłumionych na rys. 10.9 odpowiada wartości fazy początkowej $\varphi = 90°$. Innymi słowy, krzywe ciągła i kreskowana przedstawiają przebieg funkcji

$$x = A_0 e^{-\delta t} \cos \omega t$$

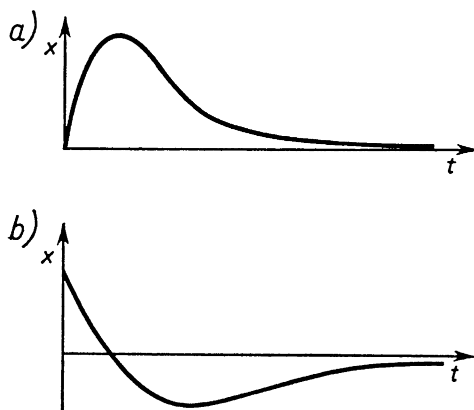
oraz

$$x = A_0 \cos \omega_0 t.$$

Pulsacje drgań tłumionych i nietłumionych (a zatem i okresy) są różne. Można wykazać, że spełniona jest zależność

$$\omega = \sqrt{\omega_0^2 - \delta^2}. \qquad (10.21)$$

Do charakteryzowania przebiegu drgań tłumionych często stosuje się tzw. logarytmiczny dekrement tłumienia oznaczany symbolem Λ. *Dekrement logarytmiczny* tłu-



Rys. 10.10

mienia jest logarytmem naturalnym stosunku dwóch amplitud odpowiadających chwilom t i $(t+T)$. Innymi słowy,

$$\Lambda = \ln \frac{A_0 e^{-\delta t}}{A_0 e^{-\delta(t+T)}} = \ln e^{\delta T} = \delta T. \qquad (10.22)$$

Uwzględniając (10.22) można równanie drgań tłumionych przepisać w postaci

$$x = A_0 e^{-\Lambda t/T} \sin(\omega t + \varphi). \qquad (10.23)$$

W ośrodkach o większych stałych tłumienia (a tym samym o większych dekremen-

tach logarytmicznych tłumienia) wygaszanie drgań jest gwałtowniejsze. Aczkolwiek teoretycznie spadek amplitudy A_0 do zera powinien nastąpić dopiero po czasie $t = \infty$, to w praktyce już po czasie skończonym, tym krótszym, im większe są δ i Λ, obserwuje się faktyczny zanik drgań.

Warto podkreślić, że badanie drgań tłumionych w określonym ośrodku pozwala wyznaczyć jego współczynnik oporu b. Z pomiaru dwóch amplitud w chwili t i w chwili $(t+T)$ można znaleźć dekrement logarytmiczny tłumienia. Uwzględniając (10.22) można wyznaczyć δ, a następnie opierając się na definicyjnej równości $b/m = 2\delta$ znaleźć współczynnik oporu b.

Powróćmy jeszcze do równania (10.21). Widać z niego, że pulsacja ω ma tylko wtedy wartość rzeczywistą, gdy wyrażenie podpierwiastkowe jest większe od zera, czyli gdy $\delta^2 < \omega_0^2$. Innymi słowy, tylko spełnienie tego warunku zapewnia powstawanie drgań tłumionych periodycznych, takich jak na rys. 10.9. W ośrodkach o silnym tłumieniu, gdy $\delta^2 \geqslant \omega_0^2$, powstaje *ruch aperiodyczny*. Dwa przykłady takiego ruchu przedstawia rys. 10.10: na rys. 10a wychylenie początkowe spada do zera bez zmiany znaku, na rys. 10b — z jednokrotną zmianą znaku.

10.5. Drgania wymuszone. Rezonans

W paragrafach 10.1 i 10.4 zajmowaliśmy się ruchem drgającym punktu materialnego w ośrodku bez tłumienia i w ośrodku z tłumieniem. W pierwszym przypadku mieliśmy do czynienia z drganiem harmonicznym nietłumionym o stałej amplitudzie A i o pulsacji własnej ω_0. W przypadku drugim (w układach o niezbyt wielkim tłumieniu) powstawały drgania tłumione o amplitudzie malejącej wykładniczo z biegiem czasu i o pulsacji ω związanej z ω_0 zależnością (10.21). Zgodnie z zapowiedzią we wstępie do niniejszego rozdziału przejdziemy obecnie do analizy drgań tzw. wymuszonych. Znów ograniczymy się do omówienia jednego z możliwych przykładów drgań wymuszonych, a mianowicie drgań powstających wtedy, gdy punkt drgający w ośrodku o stałej tłumienia δ poddany jest dodatkowo działaniu siły F sinusoidalnie zmiennej z biegiem czasu:

$$F = F_0 \sin \Omega t,$$

gdzie F_0 oznacza maksymalną wartość siły F, Ω — pulsację zmian siły F.

Uproszczoną analizę drgań wymuszonych przeprowadzimy podobnie, jak to robiliśmy w przypadku drgań tłumionych. Równanie drgań wymuszonych rozszerza się o jeden nowy wyraz w stosunku do równania drgań tłumionych (10.19) przyjmując postać

$$\frac{d^2x}{dt^2} + 2\delta \frac{dx}{dt} + \omega_0^2 x = \gamma \sin \Omega t, \tag{10.24}$$

gdzie $\gamma = F_0/m$; pozostałe symbole mają znaczenie to samo co w (10.19).

Rozwiązaniem równania (10.24) jest wyrażenie

$$x = B \sin(\Omega t - \varphi), \tag{10.25}$$

przy czym

$$B = \frac{F_0}{m\sqrt{(\omega_0^2-\Omega^2)^2+(2\delta\Omega)^2}}, \tag{10.26}$$

$$\varphi = \operatorname{arctg} \frac{2\delta\Omega}{\omega_0^2-\Omega^2}. \tag{10.27}$$

Sprawdzenie, że (10.25) jest rozwiązaniem równania różniczkowego (10.24), jest godne polecenia. Dla ułatwienia dorzucimy kilka wskazówek. Należy obliczyć pierwszą i drugą pochodną różniczkując dwukrotnie (10.25), wyniki podstawić do (10.24), otrzymane wyrażenie przekształcić stosując wzory na sinus i cosinus różnicy kątów, uporządkować skupiając wyrazy zawierające $\sin\Omega t$ i $\cos\Omega t$. Dla zapewnienia tożsamościowego spełnienia równania trzeba przyrównać współczynniki przy wspomnianym sinusie i cosinusie do zera. Z otrzymanych w ten sposób dwóch równań obliczyć B i φ.

Omówimy kilka wniosków wynikających z zależności (10.25) i (10.26). Z zależności (10.25) widać, że

1) drgania wymuszone odbywają się z pulsacją Ω siły wymuszającej,

2) drgania wymuszone mają inną fazę niż siła wymuszająca. Mówimy w takich przypadkach o przesunięciu fazowym.

Z zależności (10.26) wynika, że

3) amplituda drgań wymuszonych B — przy określonych F_0, m i δ — zależy od różnicy kwadratów pulsacji drgań własnych ω_0 i pulsacji siły wymuszającej Ω,

4) drgania wymuszone osiągają wtedy maksymalną amplitudę B_{max}, gdy wyrażenie podpierwiastkowe w mianowniku (10.26) ma minimum.

Rozpatrzmy ten przypadek bardziej szczegółowo. Gdy pulsacja siły wymuszającej jest tak dobrana, że drgania wymuszone odbywają się z maksymalną amplitudą, mówimy o *zjawisku rezonansu*. Pulsację siły wymuszającej nazywamy wtedy *pulsacją rezonansową* Ω_r. Znajdźmy warunek, jaki musi ona spełniać.

Maksimum amplitudy (czyli minimum wyrażenia podpierwiastkowego w mianowniku (10.26)) występuje wtedy, gdy

$$\frac{d[(\omega_0^2-\Omega^2)^2+(2\delta\Omega)^2]}{d\Omega} = 0.$$

Po zróżniczkowaniu otrzymujemy

$$-4\Omega(\omega_0^2-\Omega^2)+8\delta^2\Omega = 0.$$

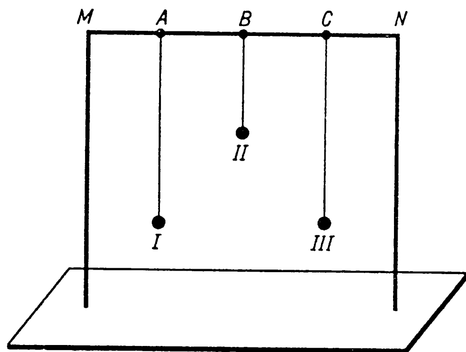
Rozwiązując to równanie względem Ω znajdujemy właśnie pulsację rezonansową Ω_r:

$$\Omega_r = \sqrt{\omega_0^2-2\delta^2}. \tag{10.28}$$

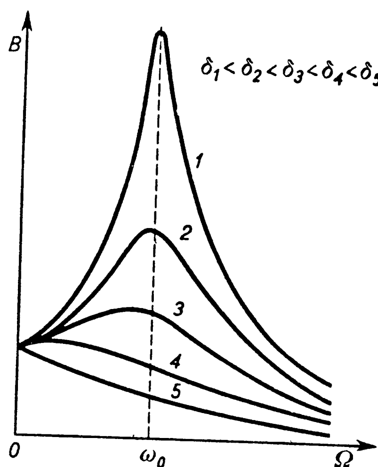
Z (10.28) widać, że *pulsacja rezonansowa jest zawsze mniejsza od pulsacji drgań własnych układu*. Różnica obu pulsacji rośnie ze wzrostem stałej tłumienia. *Dla układów o bardzo małym tłumieniu można przyjmować, że pulsacja rezonansowa siły wymuszającej jest równa pulsacji drgań własnych.* Można oczywiście też mówić w tym przypadku o równości częstotliwości siły wymuszającej i częstotliwości drgań własnych układu. (Przypominamy związek $\omega = 2\pi\nu$.) Ten warunek równości obu pulsacji lub równości obu

częstotliwości jest często — choć niezupełnie ściśle — podawany jako warunek rezonansu.

Typowym przykładem rezonansu w ośrodku o małym tłumieniu (np. w powietrzu) jest rezonans wahadeł matematycznych o jednakowej długości. W doświadczeniu przedstawionym na rys. 10.11 mamy trzy wahadła przymocowane w punktach A, B i C nici MN. Wahadła I i II mają jednakowe długości, a więc i jednakowe częstotliwości drgań własnych. Wychylając wahadło I z położenia równowagi i puszczając je swobodnie wywołujemy działanie siły okresowej na oba wahadła pozostałe. Pod działaniem tej siły wahadło III, dla którego spełniony jest warunek rezonansu, stopniowo uzyskuje dużą amplitudę, natomiast wahadło II o innej częstotliwości drgań własnych praktycznie pozostaje nieruchome.



Rys. 10.11



Rys. 10.12

Dodatkowych wiadomości o zjawisku rezonansu dostarczają tzw. *krzywe rezonansowe* (rys. 10.12). Są to krzywe przedstawiające zależność amplitudy B drgań wymuszonych od pulsacji siły wymuszającej Ω w układach o różnych stałych tłumienia δ. Krzywym *1, 2, ..., 5* odpowiadają coraz większe stałe tłumienia: $\delta_1 < \delta_2 < \delta_3 \ldots < \delta_5$. Jak widać, ośrodkom o mniejszych stałych tłumienia odpowiadają krzywe *1* i *2* o dość wyraźnie zarysowanym maksimum. Innymi słowy, w takich układach duże amplitudy mogą być osiągane tylko w wąskim przedziale wartości Ω, zawierającym pulsację rezonansową Ω_r. Krzywe *3* i *4* dotyczące ośrodków o dużym tłumieniu mają już maksima rozmyte — w szerokim przedziale wartości Ω układ drgający reaguje niewielkim wzrostem amplitudy. Wreszcie w układach o bardzo dużym tłumieniu (gdy $\delta \geqslant \omega_0/\sqrt{2}$ por. równanie (10.28)) nie obserwuje się w ogóle wzrostu amplitudy (krzywa *5*) — rezonans nie występuje.

Istnienie rezonansu — szczególnie w układach o małym tłumieniu, gdzie towarzyszy mu gwałtowny wzrost amplitudy drgań — ma swoje złe i dobre strony. Rezonans jest szeroko stosowany w celu wzmacniania drgań nie tylko zresztą mechanicznych,

lecz także akustycznych, elektrycznych itp. Czasem jednak drgania rezonansowe o dużej amplitudzie są czynnikiem bardzo niepożądanym, prowadzącym niekiedy do uszkodzenia lub zniszczenia układów drgających. Znane są m.in. przykłady zniszczenia konstrukcji mostów wskutek periodycznych pobudzeń czy to czynnikami mechanicznymi (miarowy krok żołnierzy), czy nawet czynnikami atmosferycznymi (okresowe podmuchy wiatru o częstotliwości rezonansowej), jak również przykłady występowania niebezpiecznych drgań rezonansowych obudowy maszyn, konstrukcji budowlanych itp.

10.6. Podsumowanie wiadomości o drganiach swobodnych i wymuszonych

Podane niżej zestawienie ma na celu ułatwienie usystematyzowania wiadomości związanych z niewątpliwie trudnymi zagadnieniami poruszanymi w § 10.1, 10.4 i 10.5.

Rodzaj drgań	Siły działające	Równanie ruchu	Rozwiązanie równania ruchu
1. Drgania harmoniczne bez tłumienia	$F = -m\omega_0^2 x$	$\frac{d^2x}{dt^2} + \omega_0^2 x = 0$	$x = A\sin(\omega_0 t+\varphi)$

Krótka charakterystyka: drganie sinusoidalne o stałej amplitudzie A i stałej pulsacji ω_0, zwanej pulsacją drgań własnych układu drgającego.

2. Drgania harmoniczne tłumione	$F = -m\omega_0^2 x - 2\delta m \frac{dx}{dt}$	$\frac{d^2x}{dt^2} + 2\delta \frac{dx}{dt} + \omega_0^2 x = 0$	$x = A_0 e^{-\delta t}\sin(\omega t+\varphi)$

Krótka charakterystyka: drganie o amplitudzie malejącej wykładniczo z biegiem czasu od wartości początkowej A_0 do zera i o pulsacji $\omega = \sqrt{\omega_0^2-\delta^2}$.

3. Drgania wymuszone	$F = -m\omega_0^2 x - 2\delta m \times$ $\times \frac{dx}{dt} + F_0 \sin \Omega t$	$\frac{d^2x}{dt^2} + 2\delta \frac{dx}{dt} + \omega_0^2 x =$ $= \frac{F_0 \sin \Omega t}{m}$	$x = B\sin(\Omega t-\varphi)$

Krótka charakterystyka: drganie harmoniczne o pulsacji równej pulsacji Ω siły wymuszającej, o amplitudzie $B = \frac{F_0}{m\sqrt{(\omega_0^2-\Omega^2)^2+(2\delta\Omega)^2}}$, opóźnione w fazie względem siły o kąt φ spełniający zależność $\operatorname{tg}\varphi = \frac{2\delta\Omega}{\omega_0^2-\Omega^2}$.

Przypadek szczególny drgań wymuszonych stanowią drgania rezonansowe o maksymalnej amplitudzie B_{max}, występujące przy pulsacji rezonansowej Ω_r siły wymuszającej. Pulsacja rezonansowa $\Omega_r = \sqrt{\omega_0^2-2\delta^2}$; $\Omega_r < \omega < \omega_0$.

10.7. Drgania samowzbudne. Drgania samowzbudne relaksacyjne

Omawiając różne rodzaje drgań warto chwilę uwagi poświęcić drganiom tzw. relaksacyjnym, należącym do szerszej klasy drgań samowzbudnych. Wobec dość skomplikowanej matematycznej teorii tych drgań dalszy opis będzie wyłącznie jakościowy. Celem jego będzie przedstawienie w sposób poglądowy — w oparciu o przykłady — istoty drgań samowzbudnych i drgań samowzbudnych relaksacyjnych.

Drgania samowzbudne należą do bardzo rozpowszechnionych w przyrodzie. Mogą one mieć różny charakter: mogą to być np. drgania natury mechanicznej (np. w zegarze z kołem zębatym i kotwicą) lub drgania natury elektrycznej (np. powstające w obwodzie generatora lampowego lub przerywacza magnetycznego). Drgania tego samego rodzaju są charakterystyczne np. dla pracy serca.

W każdym układzie wytwarzającym drgania samowzbudne można wyróżnić źródło energii, urządzenie regulujące dopływ energii do układu drgającego (nazwijmy je krótko zaworem) oraz układ drgający. Zawór i układ drgający sterują nawzajem swoją pracą. Ogólnie mówimy w takim przypadku o istnieniu tzw. sprzężenia zwrotnego, które zapewnia automatyczne, wewnętrzne sterowanie pracą układu. Tak więc *samowzbudny układ drgający* jest takim układem, który dzięki istnieniu źródła energii i zaworu związanego sprzężeniem zwrotnym z układem drgającym zdolny jest do wytwarzania drgań nie zanikających. Warunkiem energetycznym powstawania takich drgań jest oczywiście równość energii dostarczanej ze źródła i wydatkowanej przez układ w ciągu każdego okresu.

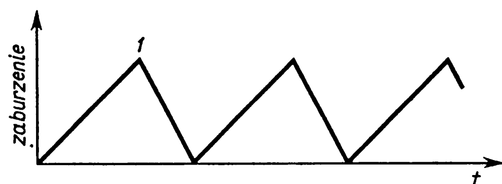
Jednym z szeroko stosowanych sposobów opisu procesu drgań jest podawanie ich zapisu graficznego, przedstawiającego zależność od czasu, np. wychylenia od położenia równowagi. Znamy już krzywe drgań sinusoidalnych, drgań zanikających, drgań aperiodycznych.

Krzywe drgań samowzbudnych mogą mieć bardzo różny przebieg, który pozwala na wprowadzenie podziału tych drgań na: 1) prawie harmoniczne i 2) relaksacyjne. Jak sama nazwa wskazuje, krzywe przebiegu drgań prawie harmonicznych są bardzo zbliżone do sinusoid. Krzywe drgań relaksacyjnych mają natomiast przebieg bardzo różny od sinusoidalnego. (W następnym rozdziale wykażemy, że drgania o takim przebiegu można uważać za wynik nałożenia bardzo dużej liczby drgań harmonicznych.)

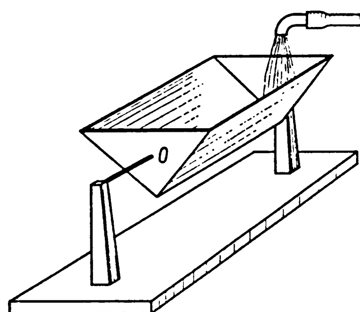
Ograniczymy się tutaj do rozpatrzenia cech jednego z możliwych przykładów drgań relaksacyjnych, a mianowicie *drgań* tzw. *piłowych*, o piłokształtnej krzywej przebiegu przedstawionej na rys. 10.13. Jak widać, w ciągu każdego okresu drgań tego typu występuje stosunkowo powolne narastanie odchylenia układu od położenia równowagi (dla zapewnienia większej ogólności ujęcia nazwijmy to krótko zaburzeniem) do wartości maksymalnej (odcinek 01), po czym następuje gwałtowny (w przypadku idealnym natychmiastowy) jego spadek do minimum. Właśnie taki okresowy, w przybliżeniu skokowy charakter zmian zaburzenia jest charakterystyczny dla drgań samowzbudnych relaksacyjnych.

Jako przykład mechanicznych drgań relaksacyjnych piłokształtnych — może dość sztuczny, lecz dobrze wprowadzający w istotne cechy tych drgań — omówimy zacho-

wanie się naczynia o przekroju np. trójkątnym, osadzonego tak, że może wykonywać drgania dookoła osi poziomej przechodzącej przez punkt O (rys. 10.14)*. Na ten układ działa stała siła, której dostarcza strumień wody padający na ściankę boczną naczynia (nie na dolną krawędź). Zbiornik stopniowo się napełnia, podnosi się jego środek ciężkości, równowaga staje się coraz bardziej chwiejna i wreszcie przy pewnej ilości wody w zbiorniku niewielka siła pochodząca od strumienia bieżącej wody wystarcza do przechylenia zbiornika, przy czym oczywiście woda się wylewa. Zbiornik wraca do położenia pierwotnego i całość zjawiska się powtarza. Odnosząc wykres na rys. 10.13 konkretnie do tego przykładu można by na osi rzędnych odkładać ilość wody w zbiorniku, na osi odciętych — czas.



Rys. 10.13



Rys. 10.14

Jeszcze prostszym przykładem procesu relaksacyjnego jest tworzenie się i odrywanie kropel wody od wylotu nieszczelnie dokręconego kranu.

Do najbardziej rozpowszechnionych drgań samowzbudnych należą drgania samowzbudne relaksacyjne, powstające dzięki istnieniu tarcia. Jako przykłady takich drgań można wymienić drgania towarzyszące wszelkiego rodzaju „skrzypom", jak skrzypienie podłóg, obuwia, tramwaju na zakręcie itp. Do tej samej kategorii zaliczyć należy drganie struny skrzypiec przy pociąganiu jej smyczkiem. W tym przypadku początkowo pod działaniem siły posuwu smyczka (dzięki sile tarcia statycznego) struna przesuwa się wraz ze smyczkiem. Ze wzrostem odkształcenia sprężystego wzrasta siła sprężysta, której wartość wreszcie przekracza wartość siły tarcia. Wtedy następuje chwilowe zerwanie kontaktu między struną i smyczkiem. Struna tracąc odkształcenie sprężyste przechodzi skokowo do nowego położenia, w którym jednak znowu styka się ze smyczkiem i opisane wyżej fazy ruchu zachodzą w tej samej kolejności. W ciągu jednego posuwu smyczka zjawisko może powtórzyć się setki razy.

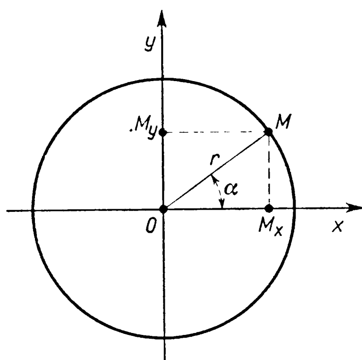
Reasumując powiemy, że drgania samowzbudne relaksacyjne mają przebieg okresowy bardzo odbiegający od sinusoidalnego, nie zanikający mimo istnienia oporów. Powstają pod działaniem siły zewnętrznej, która jednak nie jest periodycznie zmienna, jak w przypadku drgań wymuszonych rozpatrywanych w § 10.5. Przypadkiem szcze-

* Przykład zapożyczony z podręcznika A. I. Kitajgorodski, *Fizyka*, PWN, Warszawa 1963.

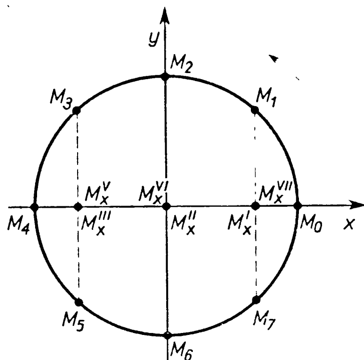
gólnym drgań relaksacyjnych są drgania piłowe. Warto też podkreślić, że wszystkie układy wykonujące drgania relaksacyjne (a nawet ogólniej — układy wykonujące drgania samowzbudne) opierają się na takich zasadach działania (sprzężenie zwrotne), które zapewniają im automatyczną powtarzalność drgań.

10.8. Składanie drgań harmonicznych

Rozważania dotyczące składania drgań harmonicznych rozpoczniemy od analizy składania drgań harmonicznych wzajemnie prostopadłych. Jak się zaraz przekonamy, przypominając wiadomości z kinematyki ruchu jednostajnego po okręgu, zagadnienie to nie jest dla nas zupełnie nowe.



Rys. 10.15



Rys. 10.16

Ruch jednostajny punktu M po okręgu jako ruch płaski odbywający się w płaszczyźnie (x, y) (rys. 10.15) opisywaliśmy równaniami

$$x = r\cos\alpha,$$

$$y = r\sin\alpha,$$

gdzie r oznacza promień koła i jest jednocześnie maksymalnym wychyleniem od położenia zerowego w ruchach wzdłuż osi x i osi y, a więc można go oznaczyć literą A i traktować jako amplitudę obu ruchów, α — kąt zakreślany przez promień wodzący, będący liniową funkcją czasu. Można go zatem przedstawić jako $\alpha = \omega t$.

Po wprowadzeniu tych nowych oznaczeń otrzymujemy

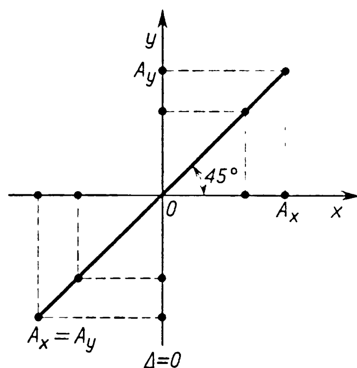
$$x = A\cos\omega t = A\sin(\omega t + \pi/2),$$

$$y = A\sin\omega t. \tag{10.29}$$

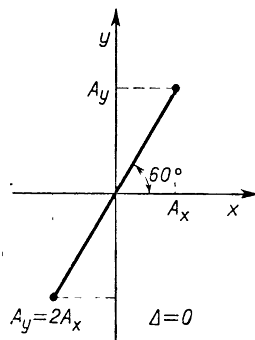
Ale wiemy już, że takie równania opisują ruch harmoniczny. Mamy więc do czynienia z dwoma ruchami harmonicznymi o jednakowej amplitudzie i pulsacji, różniącymi się w fazie o $\pi/2$, odbywającymi się w kierunkach wzajemnie prostopadłych, po dwóch wzajemnie prostopadłych średnicach. Każdy z tych ruchów harmonicznych jest ruchem rzutu punktu M na odpowiednią średnicę. Na rys. 10.16 zaznaczone są położenia

takiego rzutu na osi x co $\frac{1}{8}T$. Różnica faz wynosząca $\pi/2$ wiąże się z tym, że gdy w jednym ruchu harmonicznym punkt drgający osiąga maksymalne wychylenie od położenia równowagi, to w drugim w tej samej chwili mija on położenie równowagi. W wyniku składania opisanych dwóch ruchów harmonicznych otrzymujemy ruch jednostajny po kole (przestrzegamy jednak przed popełnianiem dość powszechnego u studentów błędu polegającego na utożsamianiu ruchu jednostajnego po kole z ruchem harmonicznym).

Po tym wstępie nawiązującym do poprzednio zdobytych wiadomości przejdziemy do bardziej systematycznego zestawienia wyników składania dwóch wzajemnie prostopadłych ruchów harmonicznych (np. wzdłuż osi współrzędnych x i y) przy różnych warunkach narzuconych okresom, amplitudom i fazom obu ruchów. Charakterystykę opisową każdego z rozważanych przypadków powiążemy z odpowiednim obrazem graficznym, pozostawiając jednak Czytelnikowi sprawdzenie przynajmniej kilku z podanych rysunków przez systematyczne sumowanie geometryczne wychyleń punktu drgającego od położenia równowagi w obu ruchach harmonicznych, np. co $\frac{1}{8}$ okresu, podobnie jak to przykładowo zrobiono na rys. 10.17.



Rys. 10.17



Rys. 10.18

Wszystkie rozpatrzone niżej przypadki oznaczone jedynką i kolejną literą alfabetu dotyczą składania ruchów harmonicznych o jednakowych okresach $T_x = T_y = T$ (a zatem i jednakowych pulsacjach $\omega_x = \omega_y = \omega$).

1 a. Amplitudy jednakowe $A_x = A_y = A$; różnica faz $\Delta = 0$.

W obu rozważanych ruchach harmonicznych punkty drgające równocześnie przechodzą przez położenia równowagi, równocześnie osiągają maksymalne wychylenia dodatnie itd. Przy odpowiednim wyborze początku rachuby czasu równania obu drgań można zapisać w postaci

$$x = A_x \sin \omega_x t = A \sin \omega t,$$

$$y = A_y \sin \omega_y t = A \sin \omega t.$$

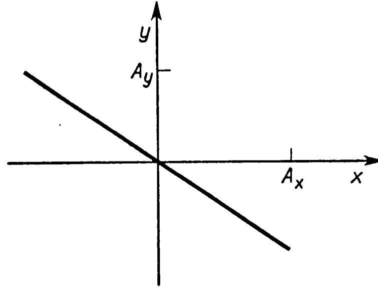
W wyniku złożenia tych drgań otrzymuje się drganie harmoniczne o amplitudzie $A_w = \sqrt{A_x^2 + A_y^2} = A\sqrt{2}$ wzdłuż linii prostej nachylonej pod kątem 45° względem osi x (rys. 10.17).

1 b. Amplitudy niejednakowe $A_x \neq A_y$; różnica faz $\Delta = 0$.

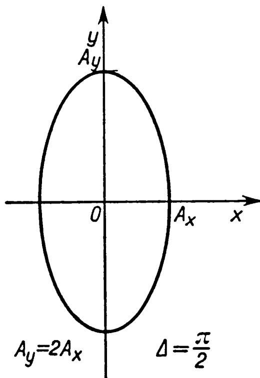
Drganie wypadkowe, podobnie jak poprzednio, jest ruchem harmonicznym o amplitudzie $A_w = \sqrt{A_x^2 + A_y^2}$, odbywającym się wzdłuż linii prostej nachylonej względem osi x pod kątem α takim, że $\operatorname{tg}\alpha = A_y/A_x$ (rys. 10.18 dla $A_y = 2A_x$).

Pytanie. Jakie są w tym przypadku równania obu drgań składowych?

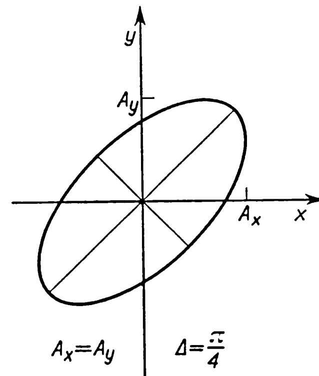
Łatwo sprawdzić, że przy różnicy faz $\Delta = \pi$ i przy amplitudach dowolnych drganie wypadkowe odbywa się również po linii prostej, tworzącej z dodatnim kierunkiem osi x kąt większy od 90° (rys. 10.19).



Rys. 10.19



Rys. 10.20



Rys. 10.21

1 c. Amplitudy jednakowe $A_x = A_y = A$; różnica faz $\Delta = \pi/2$.

Jest to rozpatrywany przez nas na początku tego paragrafu przypadek drgania wypadkowego w postaci ruchu jednostajnego po okręgu (rys. 10.16). Równania drgań składowych podane jako równanie (10.29).

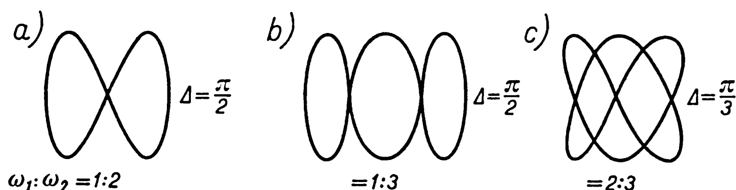
1 d. Amplitudy różne $A_x \neq A_y$; różnica faz $\Delta = \pi/2$.

Ruch wypadkowy po elipsie (rys. 10.20), której osie symetrii skierowane są wzdłuż osi x i y.

1 e. Amplitudy dowolne; różnica faz dowolna z wykluczeniem wartości 0, $\pi/2$, π, $3\pi/2$, 2π.

Ruch wypadkowy po elipsie, której osie symetrii nie są skierowane wzdłuż osi x i y (rys. 10.21 dla $A_x = A_y$ i $\Delta = \pi/4$). W zależności od wartości amplitud i różnicy faz elipsa zmienia swój kształt i położenie.

Przypadek 1 e jest oczywiście najogólniejszym przypadkiem składania drgań wzajemnie prostopadłych o jednakowych okresach. Można bowiem uważać, że otrzymywane w pozostałych przypadkach okrąg lub odcinki linii prostych są szczególnymi postaciami elips. Warto podkreślić, że we wszystkich rozważanych przypadkach okres drgania wypadkowego równa się okresowi drgań składowych.



Rys. 10.22

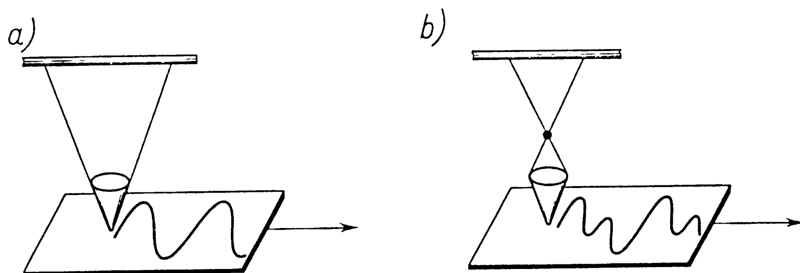
2. Ten punkt dotyczy przypadku, gdy okresy drgań składowych T_x i T_y są niejednakowe. Tym razem ruchy wypadkowe odbywają się na ogół po skomplikowanych torach krzywoliniowych, których przykłady podane są na rys. 10.22. Obok rysunku podane są warunki, jakim odpowiadają drgania składowe, a mianowicie różnica faz Δ i stosunek pulsacji.

Tory zakreślane przez punkty materialne odbywające dwa ruchy harmoniczne wzajemnie prostopadłe noszą nazwę krzywych (figur) Lissajous. Przykłady tych krzywych podane są na rys. 10.17–10.22.

Umiejętność składania ruchów harmonicznych wzajemnie prostopadłych jest bardzo przydatna przy badaniu światła spolaryzowanego (§ 26.1). Na niej też opierają się badania oscylograficzne, np. obwodów prądu zmiennego. Obrazy oscylograficzne powstają dzięki odchylaniu elektronów przez dwa wzajemnie prostopadłe zmienne pola elektryczne. Jeśli natężenia pól zmieniają się sinusoidalnie z biegiem czasu, mają te same pulsacje, lecz dowolne amplitudy i fazy, to obrazy wytwarzane przez elektrony na fluoryzującym ekranie oscyloskopu mają kształt elips, przechodzących w szczególnych przypadkach w okręgi lub odcinki linii prostych. Jeśli natomiast pulsacje obu pól są niejednakowe, to powstające obrazy oscyloskopowe mają kształt bardziej skomplikowanych krzywych Lissajous, zależnie od stosunku pulsacji i istniejącej różnicy faz natężeń obu pól.

Składanie drgań harmonicznych nie ogranicza się tylko do drgań wzajemnie prostopadłych. Regułą ogólną obowiązującą przy składaniu drgań dowolnych jest wyznaczanie wychylenia wypadkowego punktu drgającego przez sumowanie geometryczne jego wychyleń składowych w każdej chwili. Zagadnienie bardzo się upraszcza, gdy oba ruchy składowe odbywają się w tym samym kierunku (drgania równoległe), gdyż w tych warunkach sumowanie geometryczne wychyleń sprowadza się do sumowania algebraicznego (z uwzględnieniem znaków + i — wychyleń po obu stronach położe-

nia równowagi). W wyniku składania dwóch drgań harmonicznych równoległych o jednakowych pulsacjach powstaje ruch wypadkowy również harmoniczny, dla którego wykresem zależności wychylenia od czasu jest sinusoida. Jeśli jednak pulsacje obu ruchów składowych nie są jednakowe, to ruch wypadkowy nie jest już ruchem harmonicznym, aczkolwiek pozostaje ruchem okresowo się powtarzającym. Wykresem takiego ruchu nie jest już sinusoida, lecz bardziej skomplikowana krzywa.



Rys. 10.23

Omówioną zmianę charakteru drgań złożonych można pokazać w następującym doświadczeniu. Z lejka odpowiednio zawieszonego (rys. 10.23a) i wprawionego w ruch wahadłowy wysypuje się piasek na przesuwany ruchem jednostajnym w płaszczyźnie poziomej arkusz tektury. W przybliżeniu sinusoidalny rozkład piasku na tekturze przedstawia obraz wychyleń wylotu lejka od położenia równowagi w funkcji czasu. Jeśli jednak przez zawieszenie dodatkowego obciążenia, jak na rys. 10.23b zmienimy układ drgający tak, że lejek będzie równocześnie uczestniczył w dwóch ruchach drgających (własnym i układu złożonego jako całości) o różnych okresach, to układ piasku na tekturze zatraci charakter sinusoidalny. Lejek tym razem wykonuje drganie wypadkowe okresowe, lecz nieharmoniczne.

O składaniu drgań harmonicznych równoległych będziemy jeszcze mówili w następnych rozdziałach.

Pytania i zadania

1. Jak można scharakteryzować ruch harmoniczny za pomocą: a) wychylenia od położenia równowagi, b) przyspieszenia, c) siły?

2. Biorąc pod uwagę cztery kolejne ćwiartki okresu ruchu harmonicznego opisanego równaniem $s = A\sin\omega t$ określić charakter ruchu w zależności od zmian prędkości.

3. Czy słuszne jest twierdzenie, że w ruchu harmonicznym przyspieszenie jest stale skierowane do położenia równowagi?

4. Podać cechy wspólne i różnice ruchów harmonicznych opisanych równaniami $s_1 = A\sin\omega t$ i $s_2 = A\cos\omega t$.

5. Zrobić wykres zależności od czasu: a) wychylenia od położenia równowagi, b) prędkości i c) przyspieszenia w ruchach harmonicznych opisanych równaniami $s = A\sin\omega t$ i $s = A\cos\omega t$.

6. Zrobić wykres zależności energii kinetycznej i potencjalnej od wychylenia w ruchu harmonicznym.

7. Przedstawić graficznie wyniki składania dwóch ruchów harmonicznych wzajemnie prostopadłych,

o jednakowych częstotliwościach, o stosunku amplitud A_x/A_y równym kolejno 1, 2, 3 i o różnicy faz Δ równej kolejno $\pi/3$, $\pi/4$ i $\pi/6$.

8. Złożyć dwa ruchy harmoniczne wzajemnie prostopadłe o jednakowych amplitudach, o stosunku okresów $T_x/T_y = 1:2$ i o różnicy faz $\Delta = \pi/2$.

9. Scharakteryzować a) warunki powstawania drgań tłumionych, b) zmienność amplitudy tych drgań w zależności od czasu, c) zdefiniować stałą tłumienia i dekrement logarytmiczny tłumienia oraz podać związek między tymi wielkościami.

10. Jak wartość stałej tłumienia wpływa na przebieg drgań tłumionych?

11. Zestawić równania różniczkowe drgań harmonicznych nietłumionych i tłumionych.

12. Czy różnica w przebiegu drgań nietłumionych i tłumionych ogranicza się do różnicy charakteru amplitud?

13. W jakich warunkach powstają drgania wymuszone? Podać równanie różniczkowe tych drgań.

14. Scharakteryzować zjawisko rezonansu w układach o małym i dużym tłumieniu.

15. Czy częstotliwość rezonansowa jest zawsze częstotliwością drgań wymuszonych? Czy każda częstotliwość drgań wymuszonych jest częstotliwością rezonansową?

16. Scharakteryzować drgania relaksacyjne i podać ich przykłady.

17. Punkt materialny wykonuje drganie harmoniczne o amplitudzie równej A metrów. Całkowita energia tego ruchu wynosi E_c dżuli. W jakiej odległości od położenia równowagi znajduje się punkt drgający, gdy działa na niego siła równa F niutonów?

Odp. $\frac{FA^2}{2E_c}$.

18. Masa m rtęci wlana do naczynia połączonego o kształcie litery U wykonuje drgania dokoła położenia równowagi. Obliczyć okres tych drgań, jeżeli wiadomo, że przekrój ramion naczynia równa się S, a gęstość rtęci ϱ.

Odp. $2\pi\sqrt{m/2S\varrho g}$.

19. Obliczyć pulsację drgań harmonicznych nietłumionych punktu materialnego o masie $m = 10$ g, jeśli wiadomo, że amplituda drgań $A = 10$ cm, a całkowita energia ruchu $E_c = 0{,}5$ J.

Odp. 100 rad/s.

20. Wiedząc, że w pewnym ruchu harmonicznym tłumionym po okresie $T = 1$ s amplituda drgania zmniejszyła się o 0,6 amplitudy poprzedniej obliczyć stałą tłumienia oraz pulsację drgań własnych układu.

Odp. 0,91 s^{-1}, 6,3 rad/s.

21. Obliczyć pulsację rezonansową oraz amplitudę rezonansową drgań harmonicznych wymuszonych punktu materialnego o masie $m = 0{,}2$ kg, jeżeli współczynnik tłumienia $\delta = 5\ s^{-1}$, pulsacja drgań własnych $\omega_0 = 25$ rad/s i amplituda siły wymuszającej $F_0 = 14$ N.

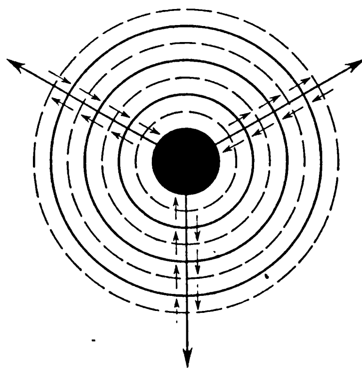
Odp. ok. 28 cm, 24 rad/s.

ROZDZIAŁ 11

Ruch falowy

11.1. Ogólna charakterystyka fal mechanicznych

Zjawisko bardzo rozpowszechnionego w przyrodzie ruchu falowego rozważane jest w wielu działach fizyki, jak to widać bezpośrednio z takich nazw jak *fale mechaniczne*, *głosowe*, *elektromagnetyczne* itp. W tym rozdziale zapoznamy się z ogólnymi własnościami fal na przykładzie fal mechanicznych, zwanych też sprężystymi, gdyż ich istnienie jest ściśle związane z właściwościami sprężystymi ośrodków, w jakich są wytwarzane.



Rys. 11.1

Niech w ośrodku modelowym doskonale sprężystym, jednorodnym, izotropowym i nieograniczonym znajduje się sprężysta kula (rys. 11.1). Jednorazowemu rozszerzeniu tej kuli bez zmiany kształtu (odkształcenie objętościowe) towarzyszy wystąpienie w jej najbliższym otoczeniu zaburzenia w postaci zgęszczenia cząstek ośrodka. Dzięki właściwościom sprężystym ośrodka zaburzenie to przekazywane jest z biegiem czasu cząstkom położonym coraz dalej od kuli, a każda z cząstek, do której doszło zaburzenie, rozpoczyna ruch drgający dokoła położenia równowagi. Kula stanowi w tym przypadku źródło pojedynczego zaburzenia w postaci zgęszczenia, czyli źródło *pojedynczego impulsu zgęszczenia* lub pojedynczej fali zgęszczenia.

Analogicznie sprawa się przedstawia w przypadku jednorazowego skurczenia się kuli. W najbliższym jej otoczeniu powstaje zaburzenie w postaci rozrzedzenia cząstek ośrodka. Dzięki sprężystości ośrodka zaburzenie to jest przekazywane cząstkom położonym coraz dalej, a każda z cząstek, do której doszło zaburzenie, rozpoczyna ruch drgający dokoła położenia równowagi. Tym razem kula stanowi źródło pojedynczego zaburzenia w postaci rozrzedzenia, tj. źródło *pojedynczego impulsu rozrzedzenia* lub pojedynczej fali rozrzedzenia.

Jeżeli zmiany objętości kuli zachodzą okresowo, to w otaczającym ośrodku sprężystym powstaje ciąg okresowo powtarzających się zgęszczeń i rozrzedzeń, czyli powstaje ciąg fal okresowych. Jak poprzednio, każde z zaburzeń dociera stopniowo do coraz dalej położonych cząstek ośrodka sprężystego zapoczątkowując ich drgania dokoła położenia równowagi. Najprostszym przypadkiem fali okresowej jest *fala harmoniczna*, powstająca wtedy, gdy źródło fali i każdy z punktów ośrodka, gdy już dotrze do niego zaburzenie, wykonuje drganie harmoniczne dokoła położenia równowagi. Większość naszych dalszych rozważań będzie dotyczyła fal harmonicznych.

Z tego, co było powiedziane wyżej, widać wyraźnie, że charakter ruchu falowego bardzo się różni od charakteru ruchu postępowego czy obrotowego. W ruchu falowym źródło fali przekazuje pewną ilość energii swemu otoczeniu. Skutkiem przekazania tej energii jest powstanie określonego zaburzenia ośrodka. To zaburzenie i związana z jego powstaniem energia przenoszą się stopniowo do coraz dalej położonych cząstek ośrodka. W ośrodku jednorodnym odbywa się to ruchem jednostajnym. Równocześnie z rozchodzeniem się zaburzenia (równoważnym — jeszcze raz to podkreślmy — transportowi energii) rozpoczynają się drgania cząstek ośrodka, do których zaburzenie już dotarło, dokoła ich położeń równowagi (drgania harmoniczne). Tak więc *ruch falowy jest związany z dwoma procesami: z transportem energii przez ośrodek od cząstki do cząstki i z ruchem drgającym poszczególnych cząstek dokoła ich położenia równowagi. Nie jest natomiast związany z ruchem materii jako całości.*

Przypomnijmy w tym miejscu znaną z życia codziennego obserwację. W pobliżu małego liścia leżącego na powierzchni stawu uderza o wodę kamień. Charakterystyczne kręgi rozchodzące się po powierzchni wody obrazują ruch zaburzenia. Drgania liścia w górę i w dół bez przesuwania się względem brzegów obrazują ruch drgający cząstek ośrodka.

W dalszych naszych rozważaniach będziemy niejednokrotnie stosowali takie pojęcia, jak promień fali i powierzchnia falowa. Przez *promień fali* rozumiemy każdy kierunek rozchodzenia się zaburzenia. Tak więc w omawianym przykładzie z kulą drgającą w ośrodku jednorodnym i izotropowym (rys. 11.1) promieniami fali są przedłużenia wszystkich promieni kuli.

Powierzchnia falowa jest to zbiór punktów ośrodka, w których zaburzenie ma tę samą fazę drgania w danej chwili. Powierzchnia falowa najdalej odsunięta od źródła, a więc stanowiąca zbiór punktów, do których w danej chwili dociera zaburzenie, nosi nazwę *czoło fali*. W omawianym przykładzie — wobec założonych właściwości ośrodka — czoło fali i pozostałe powierzchnie falowe mają kształt kul. Kilka takich powierzchni falowych widzimy na rys. 11.1 w przekroju. Linie kreskowane oznaczają powierzch-

nie falowe, odpowiadające maksymalnym rozrzedzeniom, a linie ciągłe — maksymalnym zgęszczeniom. Środkowa ciemna część rysunku odpowiada drgającej kuli. Małe strzałki o na przemian zmieniających się zwrotach, zaznaczone w pobliżu promieni fali, zwracają uwagę na to, że cząstki ośrodka w swym ruchu drgającym przesuwają się równolegle do promienia fali, w kierunku zgodnym z kierunkiem ruchu zaburzenia w zgęszczeniach, a przeciwnym do kierunku ruchu zaburzenia w rozrzedzeniach.

Na rysunku 11.1 widać, że wszystkie promienie fali są prostopadłe do powierzchni falowych. Jest to zresztą cecha ogólna promieni fal: niezależnie od kształtu powierzchni falowej prostopadła do tej powierzchni wystawiona w określonym jej punkcie wyznacza kierunek promienia fali w tym punkcie.

11.2. Podział fal

O podstawowym podziale fal ze względu na ich charakter fizyczny wspomnieliśmy już w § 11.1. Obecnie uwzględnimy inne podstawy klasyfikacji.

W zależności od *liczby wymiarów przestrzeni*, w której fala się rozchodzi, wprowadzamy podział fal na trój-, dwu- i jednowymiarowe. Przykładem fali trójwymiarowej może być fala rozchodząca się w powietrzu w otoczeniu drgającej kuli, fali dwuwymiarowej — fala rozchodząca się w postaci kręgów na powierzchni wody, fali jednowymiarowej — fala rozchodząca się wzdłuż cienkiej struny lub — w przybliżeniu — wzdłuż cienkiego węża kauczukowego.

W zależności od *kształtu powierzchni falowych* można rozróżniać fale kuliste, elipsoidalne, kołowe, eliptyczne itp. Oczywiście z wymienionych przykładowo pierwsze dwie zaliczymy równocześnie do kategorii fal trójwymiarowych, dwie następne — do dwuwymiarowych. W przypadku fali jednowymiarowej, liniowej powierzchnia falowa redukuje się do jednego punktu.

Ważną odmianą fal są *fale płaskie*. Charakteryzuje je to, że ich promienie stanowią zbiór prostych równoległych. Jeśli te promienie są rozmieszczone w przestrzeni trójwymiarowej, to powierzchnie falowe są płaszczyznami, jeśli natomiast promienie są rozmieszczone w przestrzeni dwuwymiarowej, to powierzchnie falowe redukują się do linii prostych. Warto zwrócić uwagę, że fale kuliste lub kołowe obserwowane w bardzo dużej odległości od źródła mają w przybliżeniu charakter fal płaskich.

Więcej uwagi poświęcimy podziałowi fal na poprzeczne i podłużne związanemu z kierunkiem ruchu zaburzenia w ośrodku i kierunkiem ruchu drgającego cząstek ośrodka.

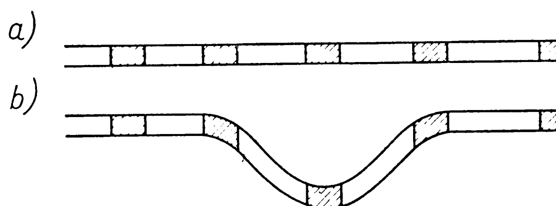
Falę poprzeczną charakteryzuje to, że w niej kierunek ruchu zaburzenia jest prostopadły do kierunku ruchu drgającego cząstek. Fale tego rodzaju powstają np. na napiętym wężu kauczukowym uderzonym prostopadle do jego długości.

W *fali podłużnej* kierunek ruchu drgającego cząstek jest równoległy do kierunku ruchu zaburzenia. Fale podłużne można wywołać na wężu kauczukowym przez jego rozciąganie i rozluźnianie. Wtedy cząstki kauczuku poruszają się w kierunku równoległym do osi podłużnej węża i w tym samym kierunku rozchodzi się wzbudzone zaburze-

nie. W omawianym wyżej przykładzie powstawania fali w ośrodku sprężystym dokoła drgającej kuli również mieliśmy do czynienia z falą podłużną.

Nasuwa się pytanie, czy fale sprężyste poprzeczne i podłużne mogą powstawać we wszystkich rodzajach ośrodków, a więc w ciałach stałych, ciekłych i gazowych.

Rozpatrzmy przesuwanie się fali poprzecznej w wężu kauczukowym. Rys. 11.2a przedstawia ośrodek w spoczynku. Podzielmy go w myśli na elementy objętości wydzielone prostopadle do osi podłużnej węża.



Rys. 11.2

Rysunek 11.2b przedstawia ten sam ośrodek w czasie przesuwania się wgłębienia. Jak widać, poszczególne elementy objętości zmieniły swój kształt, czyli uległy odkształceniu postaciowemu. Ale wiemy, że odkształcenia postaciowe mogą powstawać tylko w ciałach stałych. Tylko ciała stałe mają moduł sprężystości G nierówny zeru. *Fale poprzeczne związane z działaniem sił sprężystych mogą więc powstawać jedynie w ciałach stałych.* Aby uniknąć nieporozumień, podkreślamy, że fale powierzchniowe obserwowane tak często na powierzchni wody nie są typowymi falami sprężystymi. Przy ich powstawaniu dużą rolę odgrywa siła napięcia powierzchniowego (por. § 16.17) oraz siła ciężkości. Z tego powodu nazywamy je czasem falami grawitacyjnymi. Falami grawitacyjnymi nie będziemy się tutaj bliżej zajmowali, aczkolwiek wykorzystamy fale poprzeczne na wodzie do zilustrowania szeregu typowych zjawisk falowych.

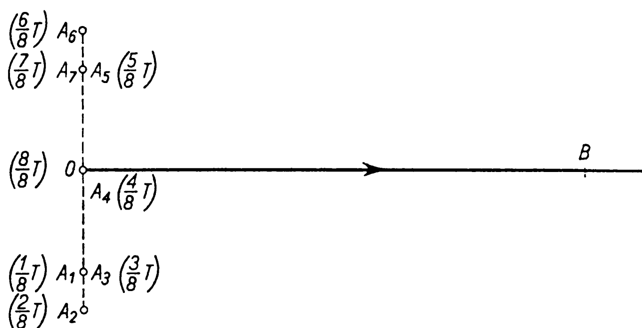
Fale podłużne polegają na rozchodzeniu się w ośrodku sprężystym zgęszczeń i rozrzedzeń, czyli związane są ze sprężystością objętościową. Sprężystość objętościową wykazują zarówno ciała stałe, jak ciekłe i gazowe. Stąd wniosek, że *fale sprężyste podłużne mogą się rozchodzić we wszystkich ciałach sprężystych, niezależnie od ich stanu skupienia.*

11.3. Mechanizm rozchodzenia się fal w ośrodkach sprężystych

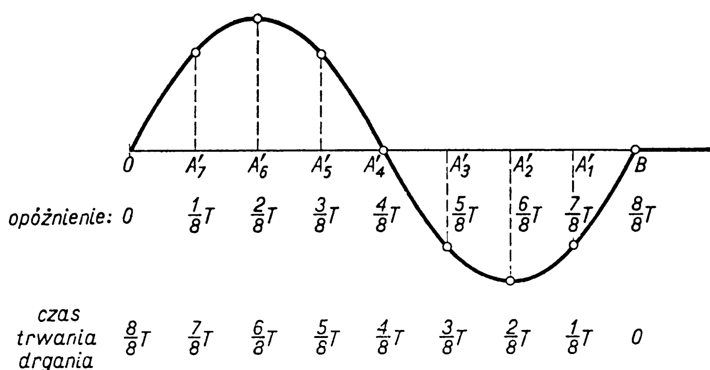
Zastanowimy się nad mechanizmem rozchodzenia się (rys. 11.3) fal harmonicznych sprężystych poprzecznych i podłużnych. Niech punkt O przedstawia źródło fali poprzecznej harmonicznej. Innymi słowy, punkt O znajdujący się w ośrodku sprężystym wykonuje drgania harmoniczne. Prześledźmy te drgania np. co $\frac{1}{8}$ okresu. W chwili $t = 0$ punkt O jest w położeniu równowagi, po $\frac{1}{8}$ okresu T mija położenie A_1, po $\frac{2}{8}T$ — położenie A_2, po $\frac{3}{8}T$ jest w A_3 ($A_3 = A_1$), po $\frac{4}{8}T$ punkt O przechodzi przez położenie równowagi A_4, po czym wychyla się w stronę przeciwną osiągając punkt A_5 po $\frac{5}{8}T$,

punkt A_6 — po $\frac{6}{8}T$, wraca do A_7 po $\frac{7}{8}T$, a po pełnym okresie znowu osiąga położenie równowagi.

Dzięki siłom sprężystym drgania wykonywane przez punkt O udzielają się sąsiednim cząstkom ośrodka. Każda z tych cząstek zostaje wprawiona w analogiczne drgania jak punkt O, ale rozpoczyna swój ruch z pewnym opóźnieniem, zależnym od jej odległości od punktu O. Przedstawmy graficznie (rys. 11.4) stan drgań punktów leżących np. na wężu kauczukowym w otoczeniu punktu O traktowanego jako źródło fali dokładnie po upływie jednego okresu od chwili rozpoczęcia ruchu. Niech OB odpowiada odległości, o jaką zaburzenie się przesunęło w ciągu jednego okresu. Wobec tego punkt



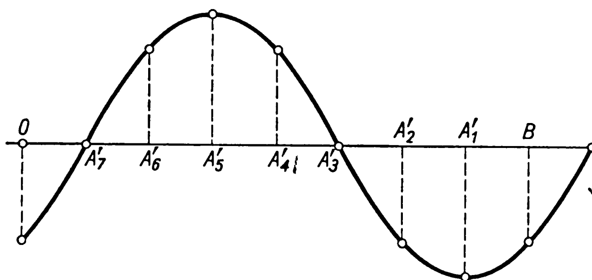
Rys. 11.3



Rys. 11.4

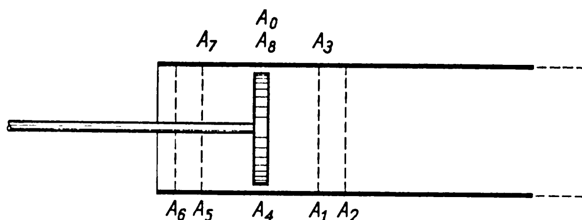
B nie ma jeszcze żadnego wychylenia od położenia równowagi. Podzielmy odległość OB np. na osiem równych części i ustalmy, w jakich ułamkach okresu odbywały się drgania punktów A'_7, A'_6, A'_5, ... itd. Do punktu A'_7 zaburzenie doszło z opóźnieniem wynoszącym $\frac{1}{8}T$, a zatem do końca okresu pozostało $\frac{7}{8}T$ i tyle wynosi czas drgania punktu A'_7. Wychylenie punktu A'_7 od położenia równowagi równa się zatem wychyleniu, jakie miał punkt O (rys. 11.3) po upływie $\frac{7}{8}T$. Wielkość OA_7 odkładamy jako rzędną punktu A'_7. Podobnie punkt A'_6 rozpoczął drgania z opóźnieniem wynoszącym $\frac{2}{8}T$, drgał więc do końca okresu w ciągu $\frac{6}{8}T$. Jego wychylenie od położenia równowagi

równa się wychyleniu, jakie miał punkt O po $\frac{6}{8}T$. To wychylenie OA_6 odkładamy jako rzędną punktu A'_6 itd. Punkt A'_1 rozpoczął drganie z opóźnieniem $\frac{7}{8}T$, zdążył uzyskać wychylenia odpowiadające $\frac{1}{8}T$, czyli jego wychylenie od położenia równowagi równa się OA_1. Punkt B spóźniony jest o $\frac{8}{8}T$, czyli po upływie jednego okresu od chwili rozpoczęcia drgań przez punkt O jest on jeszcze w położeniu równowagi, gdyż zaburzenie dopiero do niego dociera.



Rys. 11.5

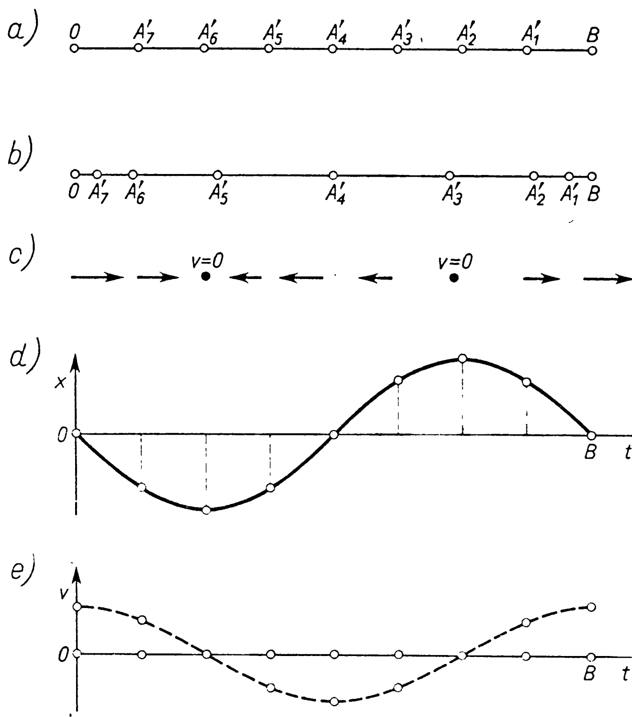
Uwzględnienie wychyleń wszystkich punktów ośrodka drgającego zawartych między O i B doprowadza do migawkowego obrazu stanu ośrodka dokładnie po upływie jednego okresu od chwili rozpoczęcia drgań przez źródło. Odpowiada on sinusoidzie z rys. 11.4. Ten stan jest jednak tylko chwilowy. Rys. 11.5 przedstawia stan drgań punktów ośrodka po upływie innego czasu, a mianowicie po czasie $(T+\frac{1}{8}T)$. Wszystkie punkty zmieniły fazy swoich drgań.



Rys. 11.6

Rozpatrzmy w podobny sposób powstawanie fali harmonicznej podłużnej. Załóżmy przykładowo, że źródłem fali jest tłok (rys. 11.6), drgający ruchem harmonicznym u wylotu długiej rury, zawierającej powietrze. Kolejne położenia tłoka co $\frac{1}{8}$ okresu odpowiadają punktom A_1, A_2, A_3 itd. Wychylenia tłoka w prawo (dodatnie) powodują zgęszczenia powietrza w rurze, wychylenia w lewo (ujemne) wywołują rozrzedzenia. Zaburzenia wzbudzone przez ruch tłoka dochodzą do coraz dalej położonych cząsteczek powietrza, ale z coraz większym opóźnieniem. Zbadajmy wychylenia cząsteczek od położeń równowagi dokładnie po upływie jednego okresu od chwili rozpoczęcia ruchu przez tłok. Niech punkt B na rys. 11.7a oznacza cząstkę ośrodka, do której dochodzi zaburzenie od źródła O po upływie 1 okresu. Dzieląc odległość OB na osiem

równych części wyodrębniamy te cząstki ośrodka, do których zaburzenie dochodzi z opóźnieniem wynoszącym $\frac{1}{8}T$, $\frac{2}{8}T$, $\frac{3}{8}T$ itd. Ich położenia równowagi przedstawiają odpowiednio punkty A_7', A_6', A_5', ..., A_1'. Każdy z tych punktów do końca okresu będzie drgał odpowiednio w ciągu $\frac{7}{8}T$, $\frac{6}{8}T$ itd. Położenia tych punktów (czyli tych cząstek ośrodka) po upływie jednego okresu od chwili rozpoczęcia ruchu przez źródło przedstawia rys. 11.7b. Wyraźnie zaznaczają się obszary zgęszczeń i rozrzedzeń cząstek. Na rysunku 11.7c zaznaczone są wektory prędkości chwilowych badanych cząstek.



Rys. 11.7

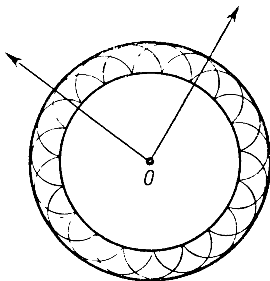
Aby otrzymać obraz fali podłużnej przypominający obraz fali poprzecznej, umawiamy się, że wielkości wychyleń od położenia równowagi odłożymy jako rzędne w stosunku do linii równowagi, uważając wychylenia w prawo za dodatnie, a w lewo za ujemne (rys. 11.7d). Stosując taką umowę otrzymujemy obrazy fali poprzecznej i fali podłużnej w postaci krzywych tego samego kształtu (sinusoidy). Badanie graficzne fali podłużnej uzupełniamy przedstawiając na rys. 11.7e rozkład prędkości poszczególnych cząstek w oparciu o rys. 11.7c, z tym jednak, że prędkości cząstek odkładamy jako rzędne. Prędkości skierowane w prawo traktujemy jako rzędne dodatnie, skierowane w lewo — jako rzędne ujemne. Ten ostatni wykres przedstawia równocześnie rozkład zmian ciśnienia liczonych względem ciśnienia równowagi w ośrodku (bez fali). W punktach O i B występuje, jak widać z rys. 11.7b, największe zgęszczenie cząstek, czyli największy wzrost ciśnienia w stosunku do ciśnienia równowagi, zaś w punkcie

A_4' panuje największe rozrzedzenie, a zatem i największe obniżenie ciśnienia w stosunku do ciśnienia równowagi.

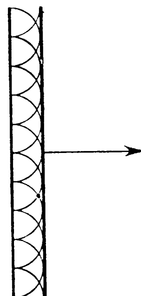
Z porównania krzywej wychyleń od położenia równowagi (rys. 11.7d) z krzywą rozkładu zmian ciśnienia (rys. 11.7e) można wyciągnąć interesujący wniosek. Podczas rozchodzenia się fali sprężystej podłużnej ekstremalne wartości zmian ciśnienia występują w tych punktach ośrodka, w których cząstki w rozważanej chwili znajdują się w położeniach równowagi. I odwrotnie, w tych punktach, gdzie występują największe wychylenia od położeń równowagi, nie ma żadnych nadwyżek lub obniżeń ciśnienia w stosunku do ciśnienia równowagi. A zatem *zmiany ciśnienia liczone w stosunku do ciśnienia równowagi są zgodne w fazie względem prędkości ruchu drgającego cząstek, a przesunięte w fazie o* $\pi/2$ *względem wychyleń cząstek od położenia równowagi.*

11.4. Zasada Huygensa. Ugięcie, odbicie i załamanie fal

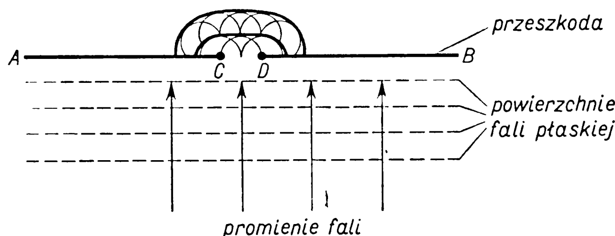
Rozchodzenie się fal o rozmaitych kształtach powierzchni falowych, jak również zjawiska ugięcia, odbicia i załamania fal można ujmować z punktu widzenia tzw. *zasady Huygensa.* Według tej zasady każdy punkt ośrodka, do którego dociera czoło fali, staje się samodzielnym źródłem wysyłającym fale kuliste cząstkowe. Powierzchnia



Rys. 11.8



Rys. 11.9



Rys. 11.10

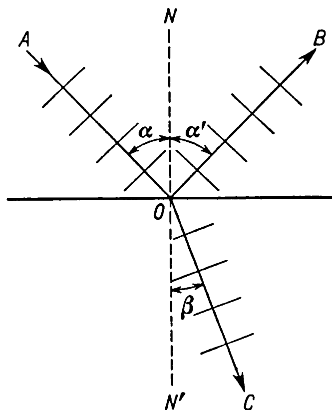
styczna do wszystkich fal kulistych cząstkowych stanowi nowe czoło fali. Rys. 11.8 i 11.9 przedstawiają właśnie w takim ujęciu rozchodzenie się fali kołowej i fali płaskiej. Rys. 11.10 przedstawia w analogiczny sposób zjawisko ugięcia fali płaskiej na małym otworze *CD* wyciętym w przeszkodzie *AB.* To samo zjawisko w odniesieniu do fal

płaskich na wodzie ilustruje rys. 11.11. Widać wyraźnie przekształcanie się fali płaskiej w falę kołową.

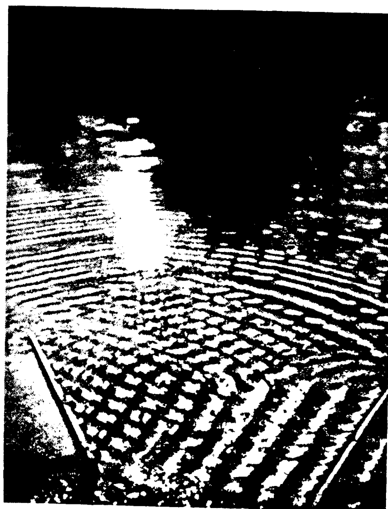
Na granicy dwóch ośrodków fala ulega zazwyczaj częściowemu odbiciu, a jeśli ośrodek drugi również jest „przezroczysty" dla danego typu fali, to równocześnie z odbiciem występuje załamanie. Oba zjawiska podlegają następującym prawom:



Rys. 11.11



Rys. 11.12



Rys. 11.13

1. *Promień fali padającej, fali odbitej i normalna wystawiona w punkcie padania leżą w jednej płaszczyźnie* (rys. 11.12).

2. *Kąt padania α równa się kątowi odbicia α'.* Kąt padania α jest to kąt zawarty między promieniem padania AO i normalną ON do powierzchni odbijającej, wystawioną w punkcie padania. *Kąt odbicia α'* jest to kąt zawarty między promieniem OB fali odbitej i normalną ON. Zaznaczone na rys. 11.12 prostopadłe do promieni przedstawiają odpowiednio powierzchnie płaskiej fali padającej i odbitej. Rys. 11.13 ilustruje odbicie fal płaskich na wodzie.

3. *Promień fali padającej, fali załamanej i normalna wystawiona w punkcie padania leżą w jednej płaszczyźnie* (rys. 11.12).

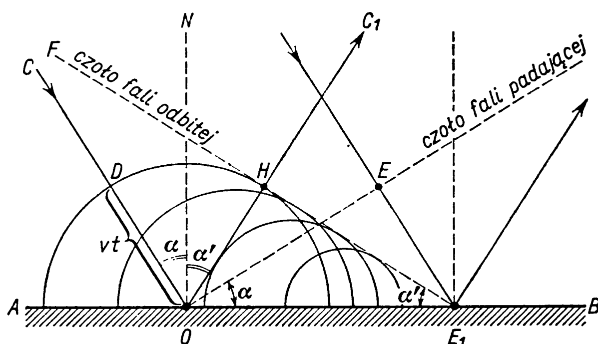
4. *Stosunek sinusa kąta padania α do sinusa kąta załamania β równa się stosunkowi prędkości rozchodzenia się fali w ośrodku pierwszym do prędkości rozchodzenia się fali w ośrodku drugim. Kąt załamania β* jest to kąt zawarty między promieniem załamanym OC i normalną ON'.

Treść ostatniego prawa można zapisać następująco:

$$\frac{\sin\alpha}{\sin\beta} = \frac{v_1}{v_2} = n\,. \tag{11.1}$$

Stałą wartość tego stosunku dla dwóch danych ośrodków i danego rodzaju fali nazywamy *współczynnikiem załamania ośrodka drugiego względem pierwszego i oznaczamy literą n.*

Przy prostopadłym padaniu promienia na powierzchnię graniczną załamanie fali nie występuje.



Rys. 11.14

Prawa odbicia i załamania fal można wyprowadzić stosując zasadę Huygensa. Niech prosta AB na rys. 11.14 przedstawia przeszkodę odbijającą. CO jest jednym z promieni fali padającej, OE stanowi część czoła fali padającej. Według zasady Huygensa punkt O staje się źródłem fali kulistej cząstkowej, z chwilą gdy do niego dociera czoło fali. Ta fala cząstkowa stopniowo z biegiem czasu powiększa swój promień, który osiąga wartość $DO = EE_1 = vt$ po czasie t, w ciągu którego czoło fali dociera od punktu E do punktu E_1. Kolejne punkty przeszkody leżące między O i E_1 stają się coraz później źródłami fal cząstkowych, więc też ich promienie są coraz mniejsze, jak to widać na rysunku. Styczna do wszystkich fal cząstkowych, przechodząca przez E_1, wyznacza czoło fali odbitej. Aby znaleźć promień fali odbitej OC_1, wystawiamy prostopadłą do jej czoła z punktu O. Uzupełniamy rysunek kreśląc normalną padania ON i zaznaczając kąt padania α i odbicia α'.

Bez trudności można wykazać równość trójkątów prostokątnych OHE_1 i OEE_1 ($OH = EE_1$ z konstrukcji), z której bezpośrednio wynika prawo odbicia

$$\alpha = \alpha'.$$

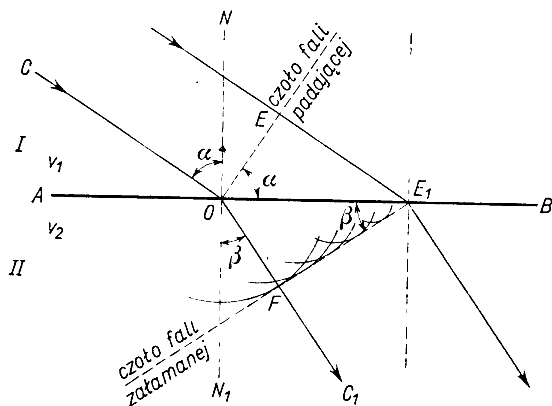
W podobny sposób dochodzimy do prawa załamania wyrażonego wzorem (11.1). Niech AB (rys. 11.15) przedstawia powierzchnię rozgraniczającą ośrodek I (prędkość fali v_1) od ośrodka II (prędkość fali v_2). Załóżmy np., że $v_1 = \frac{3}{2} v_2$. Gdy czoło fali padającej osiąga punkt O powierzchni granicznej, zaczyna on wytwarzać falę kulistą cząstkową w ośrodku II. W tym czasie, gdy w ośrodku I fala od E dociera do E_1, w ośrodku II rozchodzi się fala kulista cząstkowa o promieniu $OF = \frac{2}{3} EE_1$. Punkty pośrednie między O i E_1 wyślą do ośrodka II fale cząstkowe o odpowiednio mniejszych promieniach, jak to widać z rysunku. Styczna do wszystkich fal cząstkowych stanowi czoło fali załamanej, której promieniem wystawionym z punktu O jest OC_1. Na rysunku zaznaczone są kąty padania α i załamania β. Takie same kąty można znaleźć w trójkątach prostokątnych OEE_1 i OFE_1. Wobec tego można napisać

$$\sin\alpha = \frac{EE_1}{OE_1} = \frac{v_1 t}{OE_1}, \qquad \sin\beta = \frac{OF}{OE_1} = \frac{v_2 t}{OE_1},$$

skąd po podzieleniu stronami otrzymujemy

$$\frac{\sin\alpha}{\sin\beta} = \frac{v_1}{v_2},$$

czyli zależność (11.1).



Rys. 11.15

11.5. Równanie fali płaskiej harmonicznej

Rozpoczniemy od wyprowadzenia równania fali liniowej harmonicznej. Niech punkt O (rys. 11.16) oznacza dowolny punkt czoła takiej fali, wykonujący drgania harmoniczne. Można go uważać za źródło fali rozchodzącej się w kierunku osi Ox.

Drgania harmoniczne punktu O, odbywające się np. w kierunku osi Oy, można ogólnie wyrazić równaniem

$$y = A_0 \sin(\omega t + \varphi),$$

gdzie y oznacza wychylenie od położenia równowagi. Znaczenia pozostałych symboli znane są z rozważań poprzednich (por. § 10.1).

Początek rachuby czasu dobierzmy tak, aby faza początkowa $\varphi = 0$. Wtedy równanie na y redukuje się do wyrażenia

$$y = A_0 \sin \omega t.$$

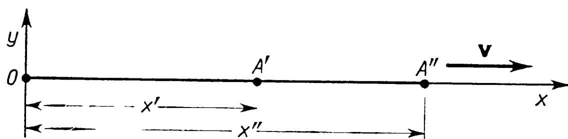
Po czasie t zaburzenie dociera do punktu A' odległego od punktu O o x'. Jest ono opóźnione względem drgania źródła o czas $\tau = x'/v$, gdzie v oznacza prędkość rozchodzenia się zaburzenia. A zatem wychylenie y' punktu A' (o współrzędnej x') od położenia równowagi, odpowiadające chwili t, wyraża się wzorem

$$y' = A_0 \sin \omega (t - \tau), \quad \text{czyli} \quad y' = A_0 \sin \omega \left(t - \frac{x'}{v}\right).$$

Podstawiając $\omega = 2\pi/T$, znajdujemy

$$y' = A_0 \sin 2\pi \left(\frac{t}{T} - \frac{x'}{vT}\right).$$

Jest to równanie fali liniowej harmonicznej.



Rys. 11.16

W punkcie A'', leżącym na tym samym promieniu w odległości x'' od źródła, wychylenie od położenia równowagi y'' w chwili t wyrazi się wzorem

$$y'' = A_0 \sin 2\pi \left(\frac{t}{T} - \frac{x''}{vT}\right).$$

Nasuwa się pytanie, jaki warunek musi spełniać odległość $(x'' - x')$, aby punkty A' i A'' były najbliższymi punktami leżącymi na tym samym promieniu fali, w których w każdej chwili wychylenia od położenia równowagi są identyczne. Innymi słowy, szukamy wyrażenia określającego tzw. *długość fali* λ.

Z okresowości funkcji sinus wynika, że y' będzie równe y'', jeżeli argumenty pod znakiem sinusa będą się różniły o całkowitą wielokrotność 2π. Punkty A' i A'' będą najbliższymi punktami o jednakowych wychyleniach, gdy

$$2\pi \left(\frac{t}{T} - \frac{x'}{vT}\right) - 2\pi \left(\frac{t}{T} - \frac{x''}{vT}\right) = 2\pi,$$

skąd

$$-\frac{x'}{vT} + \frac{x''}{vT} = 1,$$

czyli

$$x'' - x' = vT,$$

a zatem

$$\lambda = vT. \tag{11.2}$$

Ze wzoru tego wynika, że długość fali wyraża drogę, jaką zaburzenie przebywa w ciągu jednego okresu drgania źródła.

Uwzględniając (11.2) otrzymujemy wzór opisujący falę liniową w postaci

$$y = A_0 \sin 2\pi \left(\frac{t}{T} - \frac{x}{\lambda} \right). \tag{11.3}$$

Łatwo stwierdzić, że w odniesieniu do fali płaskiej otrzymamy wzór analogiczny, gdyż wszystkie punkty odległe o x od rozciągłego źródła fali płaskiej (np. w postaci drgającej listwy) będą miały w chwili t wychylenia od położenia równowagi określone równaniem (11.3).

Powróćmy raz jeszcze do równania (11.2) i spróbujmy dokładnie sprecyzować, z jaką prędkością v mamy do czynienia w równaniu definiującym długość fali. Wyobraźmy sobie obserwatora, który przesuwa się wzdłuż osi Ox razem z falą w ten sposób, że stale towarzyszy tej samej fazie drgania. Jego ruch (a równocześnie ruch wspomnianej fazy) będzie opisywała zależność $x = f(t)$, którą otrzymamy uwzględniając warunek stałości fazy, czyli

$$2\pi \left(\frac{t}{T} - \frac{x}{\lambda} \right) = \text{const},$$

skąd

$$x = \frac{\lambda}{T} t + \text{const}.$$

Różniczkując to wyrażenie względem czasu znajdziemy

$$\frac{dx}{dt} = v = \frac{\lambda}{T}.$$

Prędkość v jest zatem prędkością przesuwania się fazy w danym ośrodku. Nazywana jest *prędkością fazową* lub wprost *prędkością rozchodzenia się fali.*

11.6. Równanie różniczkowe fali płaskiej

Równanie fali (11.3) można napisać w nieco innej postaci, a mianowicie

$$y = A \sin \frac{2\pi}{\lambda} \left(t \frac{\lambda}{T} - x \right)$$

lub

$$y = A \sin \frac{2\pi}{\lambda} (vt - x). \tag{11.4}$$

Wielkość y jest funkcją dwóch zmiennych t i x, można więc obliczyć pochodne cząstkowe:

$$\frac{\partial y}{\partial t} = A \frac{2\pi}{\lambda} v \cos \frac{2\pi}{\lambda} (vt - x),$$

$$\frac{\partial^2 y}{\partial t^2} = -A \left(\frac{2\pi}{\lambda} v \right)^2 \sin \frac{2\pi}{\lambda} (vt - x) \tag{11.5}$$

oraz

$$\frac{\partial y}{\partial x} = -\frac{2\pi}{\lambda} A \cos \frac{2\pi}{\lambda} (vt - x),$$

$$\frac{\partial^2 y}{\partial x^2} = -\left(\frac{2\pi}{\lambda}\right)^2 A \sin\frac{2\pi}{\lambda}(vt-x).$$

Pomnóżmy ostatnie wyrażenie stronami przez v^2:

$$v^2\frac{\partial^2 y}{\partial x^2} = -\left(\frac{2\pi}{\lambda}v\right)^2 A \sin\frac{2\pi}{\lambda}(vt-x). \tag{11.6}$$

Z porównania (11.5) i (11.6) wynika

$$\frac{\partial^2 y}{\partial t^2} = v^2\frac{\partial^2 y}{\partial x^2}. \tag{11.7}$$

Jest to właśnie szukana ***różniczkowa postać równania fali płaskiej.***

Otrzymane równanie można uogólnić na dowolne zaburzenie ψ rozchodzące się w *przestrzeni trójwymiarowej*:

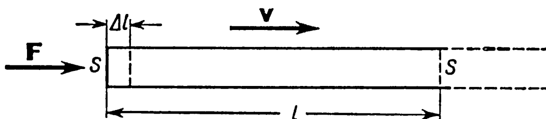
$$\frac{\partial^2 \psi}{\partial t^2} = v^2\left(\frac{\partial^2 \psi}{\partial x^2}+\frac{\partial^2 \psi}{\partial y^2}+\frac{\partial^2 \psi}{\partial z^2}\right). \tag{11.8}$$

11.7. Prędkość rozchodzenia się fali w ośrodku sprężystym

Przeanalizujmy rozchodzenie się fali podłużnej w długim sprężystym pręcie o przekroju poprzecznym S (rys. 11.17).

Niech działanie siły F w czasie Δt powoduje przesunięcie przekroju S o odległość Δl. W tym samym czasie zaburzenie wywołane działaniem siły rozchodzi się w pręcie na odległość l. Zgodnie z II zasadą dynamiki musi być spełnione równanie (2.2), czyli

$$F\Delta t = m\Delta v. \tag{11.9}$$



Rys. 11.17

Załóżmy, że wywołane działaniem siły odkształcenie pręta zawarte jest w granicach proporcjonalności, czyli, innymi słowy, zakładamy, że spełnione są warunki, w których obowiązuje prawo Hooke'a. Podstawiając zamiast naprężenia p stosunek F/S znajdujemy

$$F = SE\frac{\Delta l}{l},$$

gdzie E jest modułem Younga.

Masa cząsteczek objętych zaburzeniem równa się iloczynowi objętości Sl i gęstości materiału pręta ϱ

$$m = Sl\varrho.$$

Zmiana prędkości potrąconych cząstek pręta pozostających pierwotnie w spoczynku, a następnie przesuniętych w czasie Δt o odległość Δl wynosi

$$\Delta v = \frac{\Delta l}{\Delta t}.$$

Podstawiając te wartości do wzoru (11.9) otrzymujemy

$$SE\frac{\Delta l}{l}\Delta t = Sl\varrho\frac{\Delta l}{\Delta t}.$$

Stosunek $l/\Delta t$ wyraża prędkość rozchodzenia się zaburzenia. Oznaczmy ją symbolem v:

$$\frac{E}{v} = \varrho v,$$

$$v = \sqrt{\frac{E}{\varrho}}. \tag{11.10}$$

Jest to wzór Newtona. Jak widać ze wzoru, *prędkość rozchodzenia się fali w pręcie jest proporcjonalna do pierwiastka kwadratowego z modułu Younga i odwrotnie proporcjonalna do pierwiastka kwadratowego z gęstości.*

Prędkość nie zależy od przekroju pręta S i od wywołanego w pręcie odkształcenia, pamiętać jednak należy, że muszą być spełnione warunki stosowalności prawa Hooke'a.

Jeśli działanie siły wywołuje nie zmiany długości pręta, lecz odkształcenie objętościowe lub postaciowe, to prędkość fali wyraża się wzorami

$$v = \sqrt{\frac{K}{\varrho}} \quad \text{lub} \quad v = \sqrt{\frac{G}{\varrho}}, \tag{11.11}$$

gdzie K oznacza moduł sprężystości objętościowej, G — moduł sprężystości postaciowej.

W przypadku rozchodzenia się fal sprężystych (objętościowych) w wodzie $K = 20{,}9 \cdot 10^5 \cdot 981$ N/m^2, a więc $v = 1432$ m/s.

Przy badaniu prędkości rozchodzenia się fal sprężystych w ośrodkach gazowych należy uwzględniać rodzaj zachodzącej przemiany: czy przemiana jest *izotermiczna*, tj. zachodzi w stałej temperaturze, czy też jest *adiabatyczna*, tj. odbywa się bez wymiany ciepła z otoczeniem. W pierwszym przypadku obowiązuje prawo Boyle'a–Mariotte'a $pV = \text{const}$, w drugim — prawo Poissona $pV^\varkappa = \text{const}$ (por. § 18.6). Podczas przemian izotermicznych wartość modułu sprężystości objętościowej gazu równa się wartości ciśnienia $K = p$ (por. § 9.2). Łatwo wykażemy, że w przemianach adiabatycznych $K = \varkappa p$, gdzie $\varkappa = c_p/c_V$. W tym celu różniczkujemy równanie Poissona

$$pV^\varkappa = \text{const}$$

i otrzymujemy

$$p\varkappa V^{\varkappa-1}dV + V^\varkappa dp = 0,$$

stąd

$$dp = \frac{-\varkappa p V^{\varkappa-1}}{V^\varkappa}dV,$$

$$dp = -\varkappa p \frac{dV}{V}.$$

Znak minus przypomina, że wzrostowi ciśnienia odpowiada zmniejszanie się objętości gazu.

Porównując ostatnie wyrażenie z prawem Hooke'a

$$dp = -K \frac{dV}{V}$$

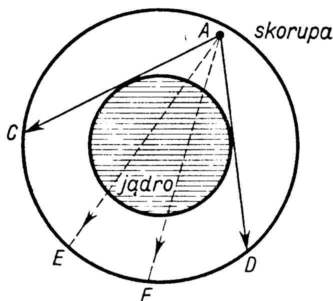
znajdujemy, że

$$K = \varkappa p. \tag{11.12}$$

Rozchodzenie się fali głosowej w gazach odpowiada w przybliżeniu przemianom adiabatycznym (czas trwania poszczególnych zgęszczeń i rozrzedzeń jest bardzo krótki). Wyznaczone doświadczalnie wartości prędkości rozchodzenia się fali głosowej w powietrzu zgodne są z wartościami obliczonymi ze wzoru

$$v = \sqrt{\frac{\varkappa p}{\varrho}}.$$

Podstawiając np. dla powietrza $\varkappa = 1{,}4$, gęstość ϱ_0 w warunkach normalnych $=$ $= 1{,}293\ \mathrm{kg/m^3}$ oraz wartość $p = 1\ \mathrm{atm} = 1013 \cdot 10^2\ \mathrm{N/m^2}$, znajdujemy prędkość rozchodzenia się głosu w powietrzu w warunkach normalnych równą 331,50 m/s.



Rys. 11.18

Wiemy już, że rozchodzenie się fal podłużnych związane jest z odkształceniami objętościowymi, a fale poprzeczne wiążą się z odkształceniami postaciowymi. Wobec różnych wartości modułów sprężystości objętościowej i postaciowej ze wzoru Newtona wynikają różne wartości prędkości rozchodzenia się obu rodzajów fal w tym samym ośrodku. Ponieważ moduł K jest większy od modułu G, więc fala podłużna rozchodzi się szybciej niż fala poprzeczna. Oba ciągi fal wzbudzone równocześnie w danym ośrodku nie biegną razem: *fala podłużna wyprzedza falę poprzeczną.*

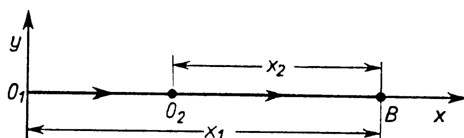
Interesujące jest rozchodzenie się takich ciągów fal w Ziemi. Gdyby Ziemia w całej swej masie była ciałem stałym, obie fale przenikałyby ją wzdłuż każdej cięciwy. Stacje sejsmiczne rejestrujące fale rozchodzące się w Ziemi wykazują inny stan rzeczy. Część środkowa Ziemi, a mianowicie kula o promieniu 0,55 całkowitego promienia, stanowi

ciekłe jądro Ziemi. Przez to ciekłe jądro mogą przenikać tylko fale podłużne. Z rys. 11.18 widać wyraźnie, że jeśli źródło trzęsienia Ziemi znajduje się w punkcie A, to obserwatoria sejsmiczne w punktach takich jak C i D odbiorą zarówno falę podłużną, jak i poprzeczną, tylko w różnych chwilach. Fala poprzeczna zostanie zarejestrowana później. Obserwatoria znajdujące się w punktach E lub F zanotują tylko odbiór fali podłużnej. Zestawiając dane różnych obserwatoriów, dotyczące czasu odbioru fal sejsmicznych, można ustalać położenie ogniska trzęsienia Ziemi, obliczać promień ciekłego jądra Ziemi oraz określać rozkład gęstości w skorupie ziemskiej.

11.8. Superpozycja fal

Załóżmy, że w pewnym ośrodku rozchodzą się równocześnie dwie fale z różnych źródeł. Fale te mogą mieć różne amplitudy i okresy. Niech obie fale wywołują niewielkie wychylenia cząstek od położenia równowagi, tzn. niech obu falom towarzyszą takie odkształcenia ośrodka, do których stosuje się prawo Hooke'a. W pewnej chwili obie fale docierają do tego samego punktu ośrodka. Każda z nich wywołuje wychylenie od położenia równowagi sobie właściwe, tzn. takie, jakie wywołałaby, gdyby drugiej fali w ogóle nie było. Wypadkowe wychylenie od położenia równowagi badanego punktu ośrodka jest równe sumie geometrycznej wychyleń składowych. Jest to tzw. *prawo składania małych wychyleń.* Gdy obie fale składowe wywołują wychylenia skierowane wzdłuż tej samej prostej, wielkości wychyleń sumują się algebraicznie. Wychylenia skierowane w tę samą stronę od położenia równowagi mają ten sam znak, wychylenia skierowane przeciwnie — mają znaki różne.

Raz jeszcze podkreślamy, że prawo składania małych wychyleń stosuje się tylko wtedy, gdy zarówno fale składowe, jak i fala wypadkowa wywołują odkształcenia spełniające prawo Hooke'a.



Rys. 11.19

Zajmiemy się szczegółowiej przypadkiem superpozycji (nakładania się) dwóch fal harmonicznych o jednakowej częstotliwości, nie wykazujących podczas rozchodzenia się żadnych przeskoków fazowych (a więc mających tzw. cechę *spójności**) wywołujących wychylenia cząstek od położenia równowagi skierowane wzdłuż tej samej prostej. Ten przypadek superpozycji nosi nazwę *zjawiska interferencji fal.*

Niech dwie fale harmoniczne o jednakowych pulsacjach rozchodzą się ze źródeł O_1 i O_2 (rys. 11.19) wzdłuż osi Ox. Oba źródła rozpoczynają drgania w tej samej chwili.

* Wielkie znaczenie wymienionego warunku spójności dla zjawiska interferencji fal będziemy mieli okazję podkreślić przy omawianiu zjawiska interferencji światła.

Wychylenia od położenia równowagi wywołane przez obie fale w dowolnie wybranym punkcie B, leżącym na osi Ox w odległości x_1 od źródła O_1 i w odległości x_2 od źródła O_2 w chwili t, wynoszą odpowiednio:

$$y_1 = A_1 \sin 2\pi \left(\frac{t}{T} - \frac{x_1}{\lambda}\right),$$

$$y_2 = A_2 \sin 2\pi \left(\frac{t}{T} - \frac{x_2}{\lambda}\right).$$

Wychylenie wypadkowe w tym punkcie jest sumą algebraiczną wychyleń składowych, czyli

$$y_w = y_1 + y_2,$$

$$y_w = A_1 \sin 2\pi \left(\frac{t}{T} - \frac{x_1}{\lambda}\right) + A_2 \sin 2\pi \left(\frac{t}{T} - \frac{x_2}{\lambda}\right).$$

Uwzględniając wzory trygonometryczne można to wyrażenie sprowadzić do postaci następującej:

$$y_w = A_w \sin(\omega t - \varphi),$$

przy czym

$$A_w^2 = A_1^2 + A_2^2 + 2A_1 A_2 \cos 2\pi \frac{x_1 - x_2}{\lambda},$$

$$\operatorname{tg} \varphi = \frac{A_1 \sin\alpha + A_2 \sin\beta}{A_1 \cos\alpha + A_2 \cos\beta},$$

$$\alpha = -2\pi \frac{x_1}{\lambda}; \qquad \beta = -2\pi \frac{x_2}{\lambda}.$$

Innymi słowy, w wyniku interferencji wymienionych fal otrzymujemy falę również o charakterze harmonicznym, lecz mającą inną niż fale składowe amplitudę A_w i fazę początkową φ.

P r z y k ł a d. Wyznaczyć parametry drgania wypadkowego w punkcie nałożenia się dwóch fal liniowych harmonicznych o jednakowych amplitudach $A_1 = A_2 = A$, jednakowych pulsacjach $\omega_1 = \omega_2 = \omega$ i fazach początkowych w punkcie spotkania φ_1 i φ_2.

Wychylenia składowe wywołane przez obie fale w punkcie spotkania wynoszą

$$y_1 = A \sin(\omega t - \varphi_1),$$

$$y_2 = A \sin(\omega t - \varphi_2).$$

Wychylenie wypadkowe y_w wynosi

$$y_w = y_1 + y_2 = A[\sin(\omega t - \varphi_1) + \sin(\omega t - \varphi_2)],$$

$$y_w = 2A \cos \frac{\varphi_2 - \varphi_1}{2} \sin\left(\omega t - \frac{\varphi_1 + \varphi_2}{2}\right).$$

Wprowadzając oznaczenia

$$A_w = 2A \cos \frac{\varphi_2 - \varphi_1}{2}, \qquad \varphi = \frac{\varphi_1 + \varphi_2}{2},$$

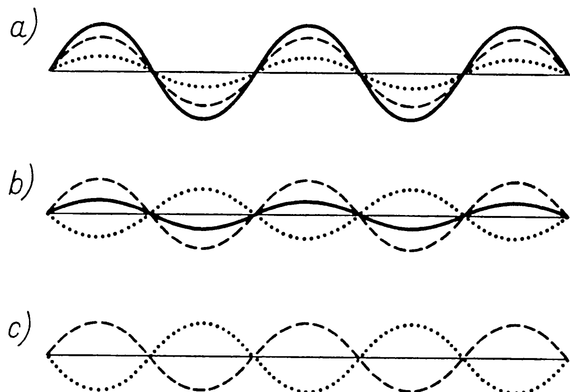
znajdujemy równanie wypadkowych drgań również typu harmonicznego

$$y_w = A_w \sin(\omega t - \varphi).$$

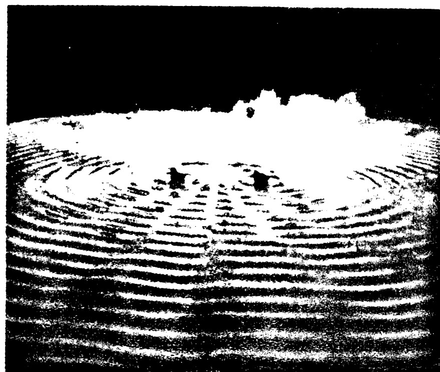
Łatwo sprawdzić, że w przypadku nakładania się fal: a) o fazach zgodnych ($\varphi_1 = \varphi_2$) występuje maksymalne wzmocnienie ($A_w = 2A$), b) o fazach przeciwnych ($\varphi_2-\varphi_1 = \pi$) — całkowite wygaszenie ($A_w = 0$).

Ogólnie biorąc, interferencja fal może prowadzić do wzmocnienia, osłabienia, a nawet zupełnego wygaszenia nakładających się fal. Rys. 11.20a, b, c przedstawiają odpowiednio interferencję dwóch fal zaznaczonych na rysunku liniami (kreskowaną i kropkowaną): a) o różnych amplitudach i zgodnych fazach, b) o różnych amplitudach i fazach przeciwnych, c) o jednakowych amplitudach i fazach przeciwnych. Grubo zaznaczone linie ciągłe przedstawiają drganie wypadkowe.

Rysunek 11.21 ilustruje interferencję fal na powierzchni wody.



Rys. 11.20



Rys. 11.21

11.9. Graficzne przedstawienie powstawania fali stojącej

Zajmiemy się obecnie graficznym przedstawieniem interferencji dwóch fal o jednakowych amplitudach A_0 i o pulsacjach ω rozchodzących się w przeciwnych kierunkach wzdłuż tego samego promienia.

Aby sprawdzić, jak zmieniają się wychylenia poszczególnych punktów podczas nakładania się takich fal, rozważymy wyniki ich spotkania, np. co $\frac{1}{8}$ okresu. Wykres odnosi się do wyodrębnionej części ośrodka, odpowiadającej jednej długości fali (rys. 11.22).

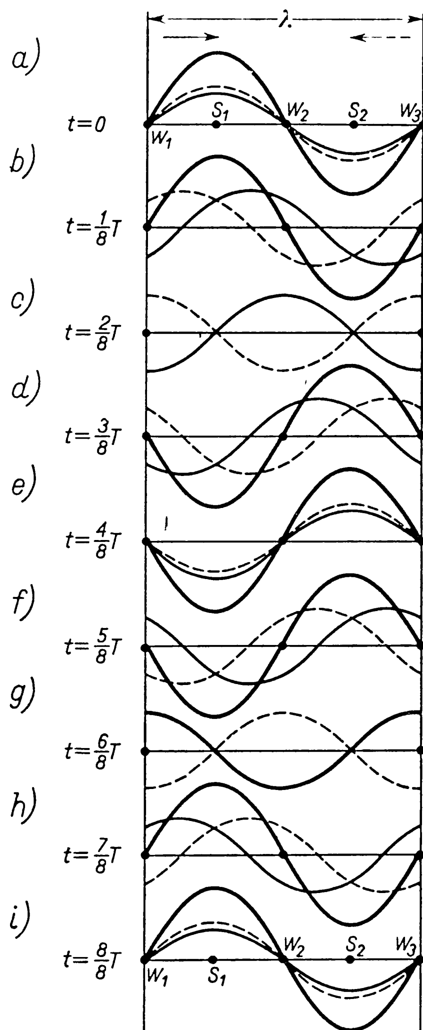
Niech w chwili $t = 0$ obie fale składowe spotykają się w wyodrębnionej części ośrodka W_1W_3 w zgodnych fazach (rys. 11.22a). Obrazem jednej fali jest sinusoida ciągła, obrazem drugiej — sinusoida kreskowana. Kierunek prędkości obu fal zaznaczony jest odpowiednio strzałką ciągłą i kreskowaną. Wobec zgodności faz i równości amplitud obu fal składowych otrzymujemy po zsumowaniu wychyleń od położeń równowagi falę wypadkową o podwojonej amplitudzie.

Rysunek 11.22b przedstawia stan tej części ośrodka po upływie $\frac{1}{8}$ okresu. Obie fale przesunęły się o $\frac{1}{8}\lambda$: fala zaznaczona linią ciągłą — w prawo, linią kreskowaną — w lewo. Znowu sumujemy wychylenia od położeń równowagi w każdym punkcie. Otrzymana fala wypadkowa ma nieco inny przebieg niż wypadkowa na rys. 11.22a. Tym

razem obie fale składowe nie spotykają się w zgodnych fazach, a więc fala wypadkowa nie ma amplitudy równej $2A_0$.

Po upływie $\frac{2}{8}T$ obie fale są przesunięte w stosunku do położeń pierwotnych o $\frac{2}{8}\lambda$, w tych samych jak poprzednio kierunkach (rys. 11.22c). Tym razem fale spotykają się w fazach przeciwnych, w tej chwili zatem wychylenia wszędzie się znoszą. Wszystkie cząstki ośrodka pozostają w tym momencie w swych położeniach równowagi.

Dalsze rysunki (11.22d, e, f, g, h, i) przedstawiają stan wyodrębnionej części ośrodka odpowiednio po $\frac{3}{8}T$, $\frac{4}{8}T$, ... i $\frac{8}{8}T$. Jak widać, kształt fali wypadkowej jest w każdej chwili inny. Dopiero po $\frac{8}{8}T$, czyli po pełnym okresie wracamy do stanu początkowego: rys. 11.22i oraz 11.22a są identyczne. Od tej chwili powtarzają się wszystkie fazy drgania.



Rys. 11.22

Spoglądając kolejno na rysunki od a do i przeanalizujmy, co się dzieje z poszczególnymi punktami badanej części ośrodka w ciągu całego okresu (a więc i w ciągu okresów następnych).

Punkt W_1 nie wykonuje żadnych drgań, jest stale w położeniu równowagi. Taki punkt nazywamy *węzłem*. Punkt S_1 odległy od W_1 o $\frac{1}{4}$ długości fali na rysunku a ma wychylenie największe, równe $2A_0$, na rysunku b ma już wychylenie nieco mniejsze, na rysunku c pozostaje w położeniu równowagi, na rysunku d ma wychylenie ujemne, które wzrasta stopniowo do maksimum na rysunku e. W następnych dwóch ósmych okresu wychylenie punktu S_1 maleje do zera, a następnie ponownie wzrasta do wartości $2A_0$. Stąd wniosek, że punkt S_1 wykonuje w ciągu jednego okresu drganie dokoła położenia równowagi osiągając co $\frac{1}{2}$ okresu wychylenie największe, równe co do wartości bezwzględnej $2A_0$.

Punkty W_2 i W_3 zachowują się w ciągu całego okresu tak jak W_1, tzn. nie wykonują żadnych drgań. Są to więc następne węzły badanej fali. Punkt S_2 zaczyna drgania w chwili $t = 0$ największym wychyleniem w dół, równym $-2A_0$, po czym stopniowo wraca do położenia równowagi, mija je w chwili $\frac{2}{8}T$, wychyla się w górę osiągając wychylenie równe $2A_0$ po $\frac{4}{8}T$, potem wraca do położenia równowagi (w chwili $t = \frac{6}{8}T$) i wychyla się ku dołowi osiągając znów wartość $-2A_0$ po upływie pełnego okresu.

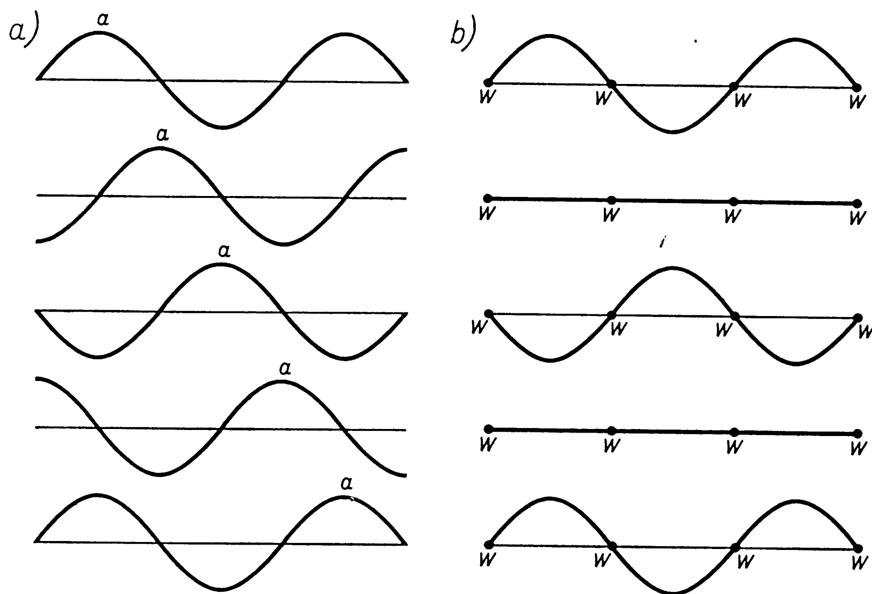
Żaden z punktów ośrodka nie osiąga wychylenia od położenia równowagi większe-

go od $2A_0$. Można więc powiedzieć, że punkty S są to takie punkty ośrodka, które mają maksymalną amplitudę drgań. Punkty te nazywamy *strzałkami. Strzałki i węzły utrzymują się w stałych punktach ośrodka. Ich istnienie w ośrodku jest charakterystyczne dla fali stojącej.*

Uwzględniając podane wyżej definicje węzłów i strzałek oraz biorąc pod uwagę charakter drgań wszystkich pozostałych punktów fali stojącej, możemy napisać, że amplituda A drgań punktów fali stojącej spełnia warunek

$$0 \leqslant A \leqslant 2A_0.$$

Równość po prawej stronie dotyczy strzałek, po lewej — węzłów, natomiast nierównością objęte są wszystkie pozostałe punkty fali stojącej. *Amplitudy drgań poszczególnych punktów są tylko i wyłącznie funkcjami ich położenia.*



Rys. 11.23

Rozpatrując jeszcze raz rys. 11.22 od a do i stwierdzamy, że każdy punkt ośrodka, np. między W_1 i W_3, wykonuje drganie dokoła położenia równowagi o amplitudzie zawartej między zerem (węzły) i wartością $2A_0$ (strzałki). Punkty leżące między W_1 i W_2 mają w każdej chwili jednakowe fazy: np. jednocześnie osiągają swe wychylenia maksymalne, jednocześnie mijają położenia równowagi itd., tzn. faza drgań tych punktów jest tylko funkcją czasu (nie zależy od ich położenia). Podobnie wszystkie punkty między W_2 i W_3 mają stale jednakowe fazy. Jednak fazy drgań punktów między W_1 i W_2 są w każdej chwili przeciwne względem faz drgań punktów między W_2 i W_3, tzn. gdy pierwsze wychylają się ku górze, to drugie w tej samej chwili przesuwają się w dół, i tak na przemian. Dzięki temu powstaje charakterystyczne dla fali stojącej „kołysanie się".

Na rysunkach 11.23a i b zaakcentowana jest *różnica między falą bieżącą i falą stojącą*. Oba rysunki przedstawiają stan drgania cząstek wyodrębnionej części ośrodka co $\frac{1}{4}$ okresu, lecz rysunek a dotyczy fali bieżącej, a rysunek b — fali stojącej. Na rysunku a widać charakterystyczne przesuwanie się fazy (np. grzbietu *a*) wzdłuż ośrodka, dla rysunku b natomiast charakterystyczna jest stałość położeń węzłów i strzałek odpowiedzialna za wspomniane „kołysanie się".

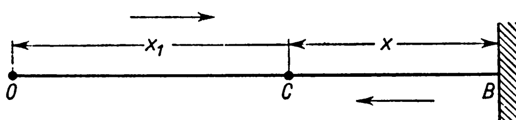
Takie kołysanie się można zaobserwować np. w przypadku fali stojącej poprzecznej na napiętym sprężystym sznurze. Tym razem fala stojąca powstaje w wyniku interferencji fali rozchodzącej się od źródła i fali odbitej od ściany, do której przymocowany jest sznur. Obie fale mają jednakowe amplitudy i pulsacje, a poruszają się w kierunkach przeciwnych. Z dalszych rozważań dotyczących fali stojącej wypłynie jeszcze dodatkowy warunek, który musi być spełniony, aby podczas interferencji fali padającej i odbitej powstała fala stojąca.

11.10. Powstawanie fali stojącej podczas interferencji fali padającej i odbitej

Rozważmy wynik interferencji fali poprzecznej, biegnącej od źródła *O*, z falą odbitą od przeszkody w punkcie *B* (rys. 11.24). Załóżmy, że podczas rozchodzenia się tych fal mamy do czynienia z niewielkimi odkształceniami sprężystymi ośrodka. Wychylenie y od położenia równowagi w dowolnie wybranym punkcie *C* równa się sumie wychyleń y_1 i y_2, wywołanych odpowiednio falę padającą i falę odbitą:

$$y_1 = A_0 \sin 2\pi\left(\frac{t}{T}-\frac{x_1}{\lambda}\right),$$

$$y_2 = A_0 \sin 2\pi\left(\frac{t}{T}-\frac{x_1}{\lambda}-\frac{2x}{\lambda}-d\right).$$



Rys. 11.24

Występujące w ostatnim wzorze wyrażenie $2x/\lambda$ wiąże się z tym, że fala odbita w porównaniu z padającą przebywa dodatkowo dwukrotnie odległość $CB = x$. Ponadto wprowadzamy parametr d, charakteryzujący warunki odbicia. Zależnie od tego, czy odbicie fali zachodzi na przeszkodzie bardziej czy mniej bezwładnej niż ośrodek, w którym fala się rozchodzi, odbicie może się odbywać z dodatkowym opóźnieniem, ze zmianą fazy $\delta = 2\pi d$. Jak się dalej okaże, podczas odbicia od ściany sztywnej, „bardziej bezwładnej", występuje zmiana fazy na przeciwną, a więc zmiana fazy $\delta = \pi$, a podczas odbicia od ściany bardziej podatnej, „mniej bezwładnej" — zmiana fazy $\delta = 0$.

Wychylenie wypadkowe y w punkcie *C*:

$$y = y_1+y_2 = A_0 \sin 2\pi\left(\frac{t}{T}-\frac{x_1}{\lambda}\right)+A_0 \sin 2\pi\left(\frac{t}{T}-\frac{x_1}{\lambda}-\frac{2x}{\lambda}-d\right).$$

Stosując wzór na sumę sinusów:

$$\sin\alpha+\sin\beta = 2\sin\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2},$$

możemy napisać

$$y = 2A_0\cos 2\pi\left(\frac{x}{\lambda}+\frac{d}{2}\right)\sin 2\pi\left(\frac{t}{T}-\frac{x_1}{\lambda}-\frac{x}{\lambda}-\frac{d}{2}\right).$$

Oznaczmy odległość od źródła fali do przeszkody odbijającej przez l:

$$l = x_1+x.$$

Wtedy

$$y = 2A_0\cos 2\pi\left(\frac{x}{\lambda}+\frac{d}{2}\right)\sin 2\pi\left(\frac{t}{T}-\frac{l}{\lambda}-\frac{d}{2}\right). \tag{11.13}$$

Jak widać, argument cosinusa jest niezależny od czasu, a zależny tylko od warunków odbicia (parametr d) oraz od odległości x wybranego punktu od ściany. Natomiast argument sinusa jest funkcją czasu.

Opierając się na analogii z ogólnym wzorem na wychylenie od położenia równowagi w ruchu harmonicznym

$$s = A\sin(\omega t+\varphi),$$

możemy przyjąć, że punkt C wykonuje drgania harmoniczne o amplitudzie

$$A = 2A_0\cos 2\pi\left(\frac{x}{\lambda}+\frac{d}{2}\right)$$

i o fazie równej

$$2\pi\left(\frac{t}{T}-\frac{l}{\lambda}-\frac{d}{2}\right).$$

Ponieważ punkt C był dowolnie wybrany, więc mamy prawo twierdzić, że podczas interferencji fali padającej i fali odbitej każdy z punktów leżących na prostej OB wykonuje ruch harmoniczny. Wzór (11.13) jest wzorem ogólnym, określającym wychylenie od położenia równowagi w chwili t dowolnego punktu leżącego na OB w odległości x od ściany odbijającej. Ponieważ amplituda jest tylko funkcją odległości x, a faza tylko funkcją czasu, więc jest to równanie fali stojącej.

Przeprowadzimy dyskusję wzoru (11.13) dla dwóch różnych warunków odbić.

1. Punkt B, w którym następuje odbicie, jest częścią masywnej ściany. Podczas odbicia fali pozostaje on nieruchomy. W tym punkcie wychylenie stale musi równać się zeru. Jest to możliwe, gdy amplituda A_B w tym punkcie jest równa zeru, czyli

$$2A_0\cos 2\pi\left(\frac{x}{\lambda}+\frac{d}{2}\right) = 0.$$

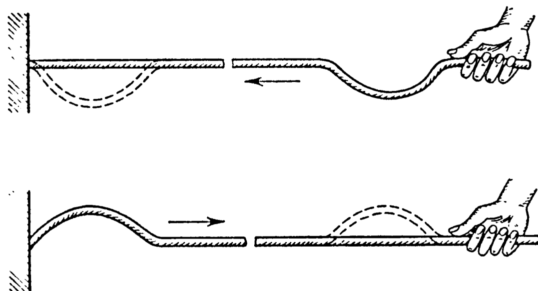
Uwzględniając, że dla punktu B odległość $x = 0$, znajdujemy

$$2\pi\frac{d}{2} = \frac{\pi}{2}, \qquad \text{czyli} \qquad d = \tfrac{1}{2}.$$

Wartości $d = \frac{1}{2}$ odpowiada zmiana fazy przy odbiciu δ wynosząca

$$\delta = 2\pi d = \pi.$$

To znaczy *odbiciu fali od ściany sztywnej towarzyszy zmiana fazy na przeciwną* (wychylenie zmienia znak). Potwierdza to doświadczenie: jeżeli do masywnej ściany dochodzi wywołana na wężu kauczukowym dolina (rys. 11.25), to po odbiciu wraca ona jako grzbiet.



Rys. 11.25

Podczas odbicia od ściany masywnej punkt leżący na samej ścianie nie wykonuje żadnych drgań — jest on węzłem fali stojącej. Ustalmy rozmieszczenie dalszych węzłów. Ze wzoru na amplitudę

$$A = 2A_0 \cos 2\pi\left(\frac{x}{\lambda}+\frac{1}{4}\right)$$

wynika, że $A = 0$, gdy

$$2\pi\left(\frac{x}{\lambda}+\frac{1}{4}\right) = (2n+1)\frac{\pi}{2},$$

$$\frac{2x}{\lambda} = n, \quad x = n\frac{\lambda}{2}, \quad \text{gdzie} \quad n = 0, 1, 2, 3, \ldots,$$

czyli

$$x' = 0, \quad x'' = \frac{\lambda}{2}, \quad x''' = 2\frac{\lambda}{2} \text{ itd.}$$

Kolejne węzły fali stojącej przypadają w odległościach równych $\lambda/2$. Zestawiamy wyniki w następującej tabelce:

Nr kolejny węzła	n	0	1	2	3
Odległość od ściany odbijającej	x	0	$\frac{\lambda}{2}$	$\frac{2\lambda}{2}$	$\frac{3\lambda}{2}$
Amplituda drgań	A	0	0	0	0

Zbadajmy, w jakich odległościach od ściany odbijającej będą leżały punkty odbywające maksymalne drgania dokoła położenia równowagi, czyli jakie jest rozmieszczenie strzałek.

Amplituda osiąga wartość maksymalną, gdy

$$\cos 2\pi\left(\frac{x}{\lambda}+\tfrac{1}{4}\right) = \pm 1,$$

czyli gdy

$$2\pi\left(\frac{x}{\lambda}+\tfrac{1}{4}\right) = n\pi,$$

$$\frac{2x}{\lambda} = n-\tfrac{1}{2},$$

$$x = n\frac{\lambda}{2}-\frac{\lambda}{4}, \quad \text{gdzie} \quad n = 1, 2, 3, \ldots$$

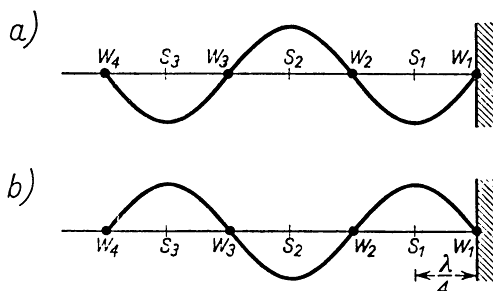
(n nie może przyjmować wartości zero, gdyż wtedy x byłoby ujemne, co odpowiadałoby punktom leżącym za ścianą odbijającą). Podstawiając kolejne wartości n znajdujemy rozmieszczenie strzałek:

$$x' = \frac{\lambda}{4}, \quad x'' = 3\frac{\lambda}{4}, \quad x''' = 5\frac{\lambda}{4} \text{ itd.}$$

Wyniki zestawiamy w następującej tabelce:

Nr kolejny strzałki n	1	2	3
Odległość od ściany odbijającej x	$\frac{\lambda}{4}$	$3\frac{\lambda}{4}$	$5\frac{\lambda}{4}$
Amplituda drgań A	$2A_0$	$2A_0$	$2A_0$

Rysunek 11.26 przedstawia stan drgań cząstek ośrodka w przypadku fali stojącej w pobliżu sztywnej ściany w chwilach różniących się o $\frac{1}{2}$ okresu.



Rys. 11.26

2. Rozważymy powstanie fali stojącej podczas odbicia od ośrodka mniej bezwładnego. Ten przypadek zachodzi np. wtedy, gdy wąż gumowy, na którym wzbudzamy falę, jest przymocowany do cienkiego sznurka i dopiero za jego pośrednictwem do ściany. Fala odbija się od sznurka, tzn. od ośrodka mniej bezwładnego. Tym razem punkt B, w którym następuje odbicie, może wykonywać drgania o maksymalnej ampli-

tudzie. Ze wzoru na amplitudę widać, że maksimum amplitudy występuje, gdy $\cos 2\pi\left(\frac{x}{\lambda}+\frac{d}{2}\right)=1$, czyli gdy argument cosinusa

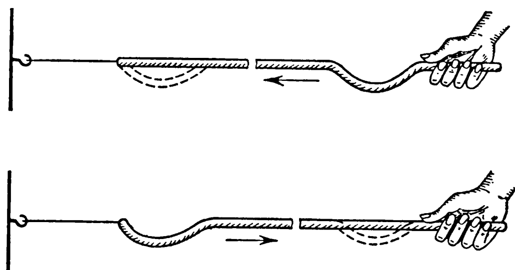
$$2\pi\left(\frac{x}{\lambda}+\frac{d}{2}\right)=0.$$

Dla punktu $B(x=0)$ wyrażenie to redukuje się do

$$2\pi\frac{d}{2}=0,$$

skąd $d=0$. A zatem podczas odbicia fali od ośrodka mniej bezwładnego nie występuje zmiana fazy, wychylenie od położenia równowagi nie zmienia swego znaku (dolina odbija się jako dolina) (rys. 11.27). Wzór na amplitudę przyjmuje postać

$$A=2A_0\cos 2\pi\frac{x}{\lambda}.$$



Rys. 11.27

Zbadajmy, jak poprzednio, rozmieszczenie węzłów i strzałek powstającej fali stojącej. W węzłach $A=0$, czyli

$$2\pi\frac{x}{\lambda}=(2n+1)\frac{\pi}{2},$$

skąd

$$x=n\frac{\lambda}{2}+\frac{\lambda}{4}, \quad \text{gdzie} \quad n=0,1,2,3,\ldots$$

Kolejne węzły występują w następujących odległościach od przeszkody odbijającej:

$$x'=\frac{\lambda}{4}, \quad x''=3\frac{\lambda}{4}, \quad x'''=5\frac{\lambda}{4} \text{ itd.}$$

Nr kolejny węzła n	0	1	2	3
Odległość od przeszkody odbijającej x	$\frac{\lambda}{4}$	$3\frac{\lambda}{4}$	$5\frac{\lambda}{4}$	$7\frac{\lambda}{4}$
Amplituda drgań A	0	0	0	0

W strzałkach występują maksymalne amplitudy, czyli spełniony jest warunek

$$2\pi\frac{x}{\lambda} = n\pi,$$

skąd

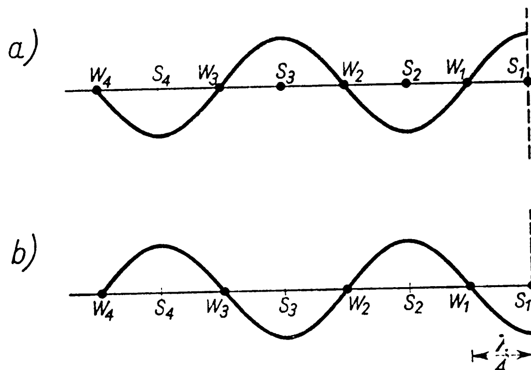
$$x = n\left(\frac{\lambda}{2}\right), \quad \text{gdzie} \quad n = 0, 1, 2, 3, \ldots$$

Kolejne strzałki powstają w następujących odległościach od przeszkody odbijającej:

$$x' = 0, \quad x'' = \frac{\lambda}{2}, \quad x''' = 2\frac{\lambda}{2} \quad \text{itd.}$$

Nr kolejny strzałki n	0	1	2	3
Odległość od przeszkody odbijającej x	0	$\frac{\lambda}{2}$	$2\frac{\lambda}{2}$	$3\frac{\lambda}{2}$
Amplituda drgań A	$2A_0$	$2A_0$	$2A_0$	$2A_0$

Rysunki 11.28a i b przedstawiają stan drgań cząstek ośrodka znajdujących się w pobliżu przeszkody odbijającej o gęstości mniejszej niż gęstość ośrodka, w którym rozchodzi się fala. Rysunki a i b odnoszą się do dwóch chwil różniących się o pół okresu.



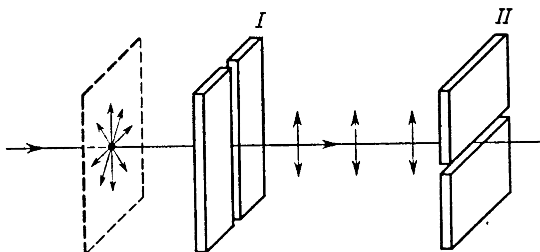
Rys. 11.28

Z tych rozważań wynika, że nie zawsze podczas interferencji fali padającej i odbitej wytworzy się fala stojąca. W celu otrzymania fali stojącej trzeba do danej odległości źródła od ściany odpowiednio dobrać częstotliwość wzbudzanej fali, aby podczas interferencji na ścianie sztywnej wytworzył się węzeł, a przy przeszkodzie mniej bezwładnej powstała strzałka fali stojącej. Jeżeli na częstotliwość wzbudzanej fali nie mamy wpływu (nie możemy jej dowolnie zmieniać), to musimy do niej dopasować odległość od źródła do ściany, aby wyżej wymienione warunki były również spełnione.

Do tego zagadnienia wrócimy jeszcze w akustyce, gdy będziemy omawiali powstawanie fal stojących w piszczałkach, drgających prętach itd.

11.11. Polaryzacja fal

Rozważmy raz jeszcze falę poprzeczną, rozchodzącą się na napiętym poziomo wężu kauczukowym. Wzbudzenie tej fali można wywołać przez uderzanie węża prostopadle do jego długości. Jeżeli uderzamy stale w tym samym kierunku, np. pionowo, to podczas rozchodzenia się fali drgania cząstek kauczuku odbywają się tylko w jednej płaszczyźnie. W omawianym przykładzie płaszczyzna drgań jest płaszczyzną pionową, określoną przez kierunek rozchodzenia się zaburzenia (oś podłużna kauczuku) i przez kierunek drgań poszczególnych cząstek (kierunek pionowy). Taką falę, w której drgania cząstek ośrodka odbywają się stale tylko w jednej płaszczyźnie zawierającej kierunek rozchodzenia się fali, nazywamy falą *spolaryzowaną liniowo.*



Rys. 11.29

Gdyby można było wzbudzić falę poprzeczną na kauczuku przez zastosowanie uderzeń prostopadłych do jego osi, ale różnokierunkowych, następujących po sobie w sposób zupełnie chaotyczny, to drgania cząstek kauczuku nie odbywałyby się stale w jednej płaszczyźnie zawierającej promień fali. Byłaby to fala niespolaryzowana. Spolaryzowanie tej fali, czyli sprowadzenie drgań do jednej płaszczyzny zawierającej kierunek rozchodzenia się fali, można osiągnąć bardzo prosto. Wystarczy ustawić dwie równoległe listwy drewniane I tak, aby wąż kauczukowy znalazł się w szczelinie między nimi (rys. 11.29). Do listew dochodzi fala niespolaryzowana. Poza listwy przechodzą bez żadnej zmiany te drgania, które się odbywają w płaszczyźnie równoległej do listew, natomiast całkowicie zostają zatrzymane (wygaszone) te drgania, które odbywają się w płaszczyźnie prostopadłej do listew. Drgania ukośne względem płaszczyzny listew zostają rozłożone na dwie składowe. Składowa równoległa do listew zostaje przepuszczona, składowa prostopadła zostaje zatrzymana. Poza listwami rozchodzi się fala liniowo spolaryzowana. Czynnik wywołujący polaryzację nazywamy *polaryzatorem.* W opisanym doświadczeniu para listew I stanowi polaryzator.

Obrót polaryzatora dokoła kierunku rozchodzenia się fali jako osi nie wpływa na amplitudę drgań przepuszczanych, jeśli tylko w fali niespolaryzowanej drgania o jednakowych amplitudach odbywały się równomiernie we wszystkich kierunkach.

Ustawienie drugiej pary listew równolegle do pierwszej nie ma żadnego wpływu na dalszy bieg fali spolaryzowanej liniowo. Wystarczy jednak drugą parę zacząć obracać dokoła kierunku rozchodzenia się fali jako osi, by wywołać zmiany amplitudy

przepuszczonego drgania oraz zmiany płaszczyzny, w której te drgania się odbywają. Znów należy uwzględnić rozkład drgania spolaryzowanego na dwie składowe: składową równoległą do drugiej pary listew i składową do niej prostopadłą. Składowa równoległa zostaje przepuszczona, składowa prostopadła — zatrzymana. Jeśli drugą parę listew skręcimy względem pierwszej o 90° (czyli skrzyżujemy obie pary), to w ogóle nie otrzymamy drgania przepuszczonego poza drugą parę listew. Stąd wniosek, że obrót pary listew dokoła kierunku rozchodzenia się fali spolaryzowanej liniowo daje inne wyniki niż obrót takiej samej pary w odniesieniu do fali niespolaryzowanej. Innymi słowy, obracana para listew może służyć do zbadania, czy dochodząca do niej fala jest spolaryzowana liniowo. Taką parę listew nazywamy *analizatorem*. Jeśli przy pewnym położeniu analizatora fala zostaje całkowicie wygaszona, oznacza to, że fala dochodząca do listew była liniowo spolaryzowana.

Gdybyśmy na wężu kauczukowym wywołali falę podłużną, to zastosowanie dowolnej liczby różnie ustawionych listew nie miałoby żadnego wpływu na rozchodzenie się fali. *Zjawisko polaryzacji występuje tylko w dziedzinie fal poprzecznych.* Stwierdzenie istnienia zjawiska polaryzacji może rozstrzygnąć wątpliwości co do charakteru fali. Tak np. wykrycie zjawiska polaryzacji światła zdecydowało o zaliczeniu fal świetlnych do fal poprzecznych.

Na tym kończymy krótką charakterystykę zjawiska polaryzacji. Do bardziej szczegółowej analizy tego zjawiska, obejmującej obok polaryzacji liniowej także kołową i eliptyczną, powrócimy jeszcze w optyce.

W tym rozdziale nie będziemy się też zajmowali zmianą częstotliwości fal odbieranych w stosunku do częstotliwości fal nadawanych, występującą wtedy, gdy źródło fali lub obserwator są w ruchu względem ośrodka, w którym rozchodzi się fala. Zjawisko to, zwane zjawiskiem Dopplera, rozpatrzymy w rozdziale poświęconym akustyce.

11.12. Fale okresowe nieharmoniczne

Nasuwa się pytanie, czy ograniczając się do rozważania fal związanych z drganiami harmonicznymi źródeł bardzo zwęziliśmy zakres badań? Okazuje się, że nie, gdyż można udowodnić, że *każde drganie okresowe nieharmoniczne można przedstawić jako sumę n drgań składowych harmonicznych.* Odpowiada to znanemu z matematyki rozkładowi funkcji okresowej na tzw. *szereg Fouriera,* który może być szeregiem nieskończonym.

Niech y rozumiane jako funkcja czasu oznacza dowolne drganie okresowe, a funkcje $y', y'', y''', \ldots, y^{(n)}$ rozumiane również jako funkcje czasu przedstawiają następujące drgania harmoniczne:

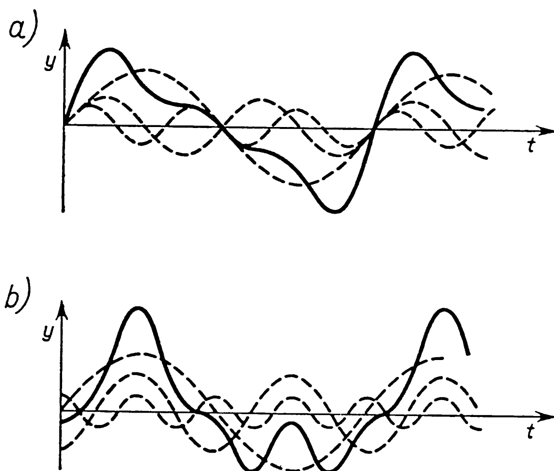
$$\begin{aligned} y' &= A_0'\sin(\omega t+\varphi'),\\ y'' &= A_0''\sin(2\omega t+\varphi''),\\ y''' &= A_0'''\sin(3\omega t+\varphi'''),\\ &\cdots\cdots\cdots\\ y^{(n)} &= A_0^{(n)}\sin(n\omega t+\varphi^{(n)}).\end{aligned}$$

Wtedy zgodnie z twierdzeniem Fouriera

$$y = A_0' \sin(\omega t + \varphi') + A_0'' \sin(2\omega t + \varphi'') + \ldots + A_0^{(n)} \sin(n\omega t + \varphi^{(n)}).$$

Jak widać z tego równania, poszczególne drgania składowe różnią się amplitudami, pulsacjami i fazami początkowymi. Pulsacje nie zmieniają się jednak w sposób dowolny. Pierwsze drganie składowe, zwane *podstawowym*, ma pulsację najmniejszą, równą ω. Dalsze drgania składowe, zwane *drganiami harmonicznymi wyższymi*, mają pulsacje 2ω, 3ω itd. Innymi słowy, *pulsacje drgań wyższych harmonicznych są całkowitymi wielokrotnościami pulsacji drgania podstawowego.*

Powiedzieliśmy, że szereg Fouriera może być szeregiem nieskończonym. Zwykle jednak amplitudy wyższych drgań harmonicznych dość szybko maleją ze wzrostem n, tak że w praktyce rozkład dowolnego drgania okresowego sprowadza się do skończonej, a nawet niewielkiej liczby drgań harmonicznych.



Rys. 11.30

Rozważania nasze dotyczące drgań harmonicznych moglibyśmy więc uogólnić na dowolne drganie okresowe przez uwzględnienie rozkładu na szereg Fouriera i przeprowadzenie sumowania w odniesieniu do odpowiedniej liczby składników tego szeregu.

Ilustracją rozkładu drgań okresowych na drgania harmoniczne jest rys. 11.30. Na rys. 11.30a drganie okresowe nakreślone grubą linią ciągłą jest sumą trzech drgań harmonicznych o różnych amplitudach, o zgodnych fazach początkowych i o częstotliwościach pozostających w stosunku jak 1:2:3. Na rys. 11.30b z tych samych drgań harmonicznych, tylko przesuniętych w fazie, powstaje zupełnie inne drganie okresowe.

11.13. Strumień energii i natężenie fali. Ciśnienie i opór akustyczny

Jak już wiemy, rozchodzenie się fali polega na przenoszeniu się energii (a nie materii) do cząstek coraz dalej położonych od źródła. Można zatem wprowadzić nową wielkość fizyczną, a mianowicie *strumień energii fali przechodzący przez daną powierzch-*

nię, będący miarą ilości energii przenoszonej przez tę powierzchnię w jednostce czasu. Strumień energii fali wyraża się zatem w jednostkach mocy, czyli w układzie SI w watach.

Jeśli strumień energii związany jest z powierzchnią prostopadłą do promienia fali (a więc stanowiącą część powierzchni falowej), to dzieląc strumień energii przez pole tej powierzchni znajdujemy nową wielkość, a mianowicie natężenie fali*. *Natężenie fali* wyraża zatem liczbowo ilość energii przenoszonej w jednostce czasu przez jednostkę powierzchni ustawionej prostopadle do kierunku rozchodzenia się fali. W układzie SI natężenie fali wyraża się w W/m^2.

Natężenie fali I można powiązać z inną wielkością, a mianowicie z gęstością energii fali. Przez *gęstość energii fali w* rozumiemy ilość energii przypadającą na jednostkę objętości ośrodka, w którym rozchodzi się fala. W fali płaskiej rozchodzącej się w nieograniczonym, nie absorbującym energii ośrodku gęstość energii fali jest stała. Między natężeniem fali i gęstością energii fali, jak łatwo się przekonać, zachodzi związek

$$I = wv, \tag{11.14}$$

gdzie v oznacza prędkość rozchodzenia się fali. Wystarczy wyobrazić sobie walec oparty podstawą na powierzchni falowej fali płaskiej, o wysokości równej liczbowo prędkości fali v i jednostkowym polu przekroju. Ilość energii zawartej w takim walcu jest właśnie liczbowo równa vw, a równocześnie przedstawia ilość energii przenoszonej w jednostce czasu przez jednostkę powierzchni prostopadłej do kierunku rozchodzenia się fali. Jest zatem równa natężeniu fali.

Energia każdego elementu objętości ośrodka, w którym rozchodzi się fala harmoniczna, wiąże się z ruchem drgającym cząstek dokoła ich położeń równowagi. Można ją zatem powiązać z takimi parametrami, jak amplituda prędkości ruchu drgającego cząstek i amplituda ich wychyleń od położenia równowagi. Rozważając tę sprawę bardziej szczegółowo dla fali podłużnej (płaskiej lub kulistej rozważanej w dużej odległości od źródła) powiążemy tę energię z jeszcze jednym parametrem, a mianowicie z amplitudą dodatkowego ciśnienia, powstającego w ośrodku w związku z wytwarzaniem się w nim charakterystycznych dla fali podłużnej zgęszczeń i rozrzedzeń. Rozważania nasze odniesiemy do gęstości energii fali.

Gęstość energii fali, czyli całkowita energia przypadająca na jednostkę objętości jest sumą energii wszystkich cząstek w tej objętości, wykonujących ruchy harmoniczne dokoła położeń równowagi. Ale z teorii ruchu harmonicznego (por. § 10.1) wiadomo, że energia całkowita każdej drgającej cząstki, niezależnie od fazy jej drgania, jest stała i równa albo maksymalnej energii kinetycznej, albo maksymalnej energii potencjalnej. Wyrażając gęstość energii jako sumę maksymalnych energii kinetycznych wszystkich cząstek zawartych w jednostce objętości (całkowita ich masa równa się liczbowo gęstości ϱ) otrzymamy

$$w = \frac{\varrho u_0^2}{2}, \tag{11.15}$$

* Stosując konsekwentnie terminologię wprowadzoną w § 9.5 należałoby tę wielkość nazywać gęstością strumienia energii fali.

gdzie ϱ oznacza gęstość ośrodka, u_0 — amplitudę prędkości ruchu harmonicznego równą, zgodnie z § 10.1, iloczynowi amplitudy drgań A i pulsacji ω, czyli

$$u_0 = A\omega. \tag{11.16}$$

Uwzględniając tę ostatnią zależność znajdziemy drugie wyrażenie na gęstość energii fali płaskiej:

$$w = \frac{\varrho A^2\omega^2}{2}. \tag{11.17}$$

Bez szczegółowych rozważań podamy trzecie wyrażenie na gęstość energii fali płaskiej podłużnej:

$$w = \frac{p_0^2}{2v^2\varrho}, \tag{11.18}$$

gdzie p_0 oznacza amplitudę omówionego wyżej dodatkowego ciśnienia.

Z porównania wyrażeń (11.18) i (11.15) otrzymujemy

$$\frac{p_0^2}{v^2\varrho} = \varrho u_0^2, \tag{11.19}$$

$$p_0 = \varrho v u_0. \tag{11.20}$$

Rozważając rys. 11.7 doszliśmy do wniosku, że zmiany ciśnienia dodatkowego p są przeciwne w fazie względem wychylenia od położenia równowagi. Są zatem zgodne w fazie z prędkością ruchu drgającego cząstek. Znając amplitudę tego ciśnienia (wzór (11.20)) można więc napisać

$$p = \varrho v u_0 \cos(\omega t + \varphi). \tag{11.21}$$

Wszystkie symbole mają to samo znaczenie, co poprzednio.

Rozważania, które w tej chwili przeprowadzamy, mają specjalne znaczenie w akustyce i dlatego — mimo że są one słuszne ogólnie dla fal podłużnych mechanicznych — wprowadzono następujące nazwy nawiązujące do akustyki: *ciśnienie akustyczne* — dodatkowe ciśnienie p ponad ciśnienie równowagi ośrodka (wyraża je wzór (11.21)), *amplituda ciśnienia akustycznego* — wielkość p_0, *opór akustyczny* — iloczyn ϱv.

Ta ostatnia nazwa tłumaczy się tym, że jak wynika z (11.20), im większy jest iloczyn ϱv, tym mniejsza jest amplituda prędkości cząstek ośrodka przy tej samej amplitudzie ciśnienia akustycznego. Ta sama zależność obowiązuje też dla każdej chwilowej wartości prędkości cząstek i ciśnienia akustycznego wobec zgodności faz obu wielkości.

Uwzględniając wzory (11.14) oraz (11.15), (11.17) i (11.18) otrzymamy następujące wyrażenia na natężenie fali:

$$I = \tfrac{1}{2}\varrho u_0^2 v, \tag{11.22}$$

$$I = \tfrac{1}{2}\varrho A^2\omega^2 v, \tag{11.23}$$

$$I = \frac{1}{2}\frac{p_0^2}{\varrho v}. \tag{11.24}$$

Są to wzory wiążące natężenie fali płaskiej podłużnej z amplitudą prędkości, amplitudą wychylenia oraz amplitudą ciśnienia akustycznego.

Pytania i zadania

1. Wymienić kilka możliwych klasyfikacji fal.

2. Zdefiniować podstawowe pojęcia związane z ruchem falowym, jak: powierzchnia fali, czoło fali, promień fali, prędkość fazowa fali, długość fali.

3. Omówić szczegółowo mechanizm rozchodzenia się fal: a) poprzecznych, b) podłużnych w ośrodku sprężystym.

4. Podać treść zasady Huygensa. Wykorzystać tę zasadę do: a) wyjaśnienia takich zjawisk, jak rozchodzenie się fali płaskiej, odbicie fali płaskiej od przeszkody kulistej, ugięcie fali na małym otworze w przeszkodzie, b) wyprowadzenia praw odbicia i załamania fal.

5. Omówić wyprowadzenie dwóch postaci równania fali płaskiej harmonicznej.

6. Podać plan wyprowadzenia wzoru Newtona na prędkość fali w ośrodku sprężystym. Wymienić podstawowe założenia.

7. Omówić warunki występowania zjawiska superpozycji fal. Na jakim prawie opiera się wyznaczanie wypadkowego wychylenia od położenia równowagi?

8. Jak można określić zjawisko interferencji fal?

9. Podać dokładną charakterystykę fali stojącej. Jakie warunki muszą być spełnione, aby podczas interferencji dwóch fal bieżących wytworzyła się fala stojąca?

10. Zrobić szczegółowy plan rozważań matematycznych związanych z wytwarzaniem się fali stojącej podczas odbicia od masywnej ściany?

11. Czy podczas odbicia fali od ośrodka bardziej bezwładnego i mniej bezwładnego zachodzą jednakowe przeskoki fazowe?

12. Podać ogólną charakterystykę zjawiska polaryzacji fal. Czy możliwa jest polaryzacja w dziedzinie fal podłużnych?

13. Jakie twierdzenie łączy się z rozkładem dowolnego drgania okresowego na drgania harmoniczne?

14. Co rozumiemy przez strumień energii fali przechodzący przez daną powierzchnię? Przy jakiej okazji była już w kursie fizyki mowa o strumieniu pewnej wielkości?

15. Co rozumiemy przez natężenie fali w danym punkcie ośrodka?

16. Statek płynący po jeziorze wywołał falę, która doszła do brzegu po upływie 10 minut. Odległość między sąsiednimi grzbietami fali wynosiła 1,5 m, a czas między kolejnymi uderzeniami fali o brzeg wynosił 2 s. Znaleźć odległość między statkiem i brzegiem.

Odp. 450 m.

17. Od źródła drgań harmonicznych rozchodzi się fala liniowa. Znaleźć wychylenie od położenia równowagi punktu oddalonego od źródła o $\frac{3}{4}$ długości fali po upływie 0,9 okresu od chwili rozpoczęcia drgań przez źródło. Amplituda drgań źródła wynosi 10 cm. Absorpcję energii w ośrodku zaniedbujemy.

Odp. 8,1 cm.

18. Równanie ruchu źródła fali harmonicznej dane jest w postaci

$$y = 10\sin 0{,}5t \text{ cm}.$$

Znaleźć a) równanie fali, jeśli wiadomo, że prędkość fali wynosi $v = 300$ m/s, b) równanie drgań punktu odległego o 600 m od źródła, c) równanie drgań punktów fali po upływie czasu $t = 4$ s od chwili rozpoczęcia drgań źródła.

19. Prędkość rozchodzenia się fal podłużnych w stali $v_1 = 5100$ m/s. Jaka jest prędkość fal poprzecznych w stali, jeśli liczba Poissona wynosi dla stali 3,1.

Odp. 3111 m/s.

20. Obliczyć stałą $\varkappa$ dla helu, jeśli wiadomo, że prędkość głosu w helu w warunkach normalnych $v = 940$ m/s. Masa atomowa helu równa się 4.

Odp. 1,56.

21. W wyniku interferencji dwóch fal o częstotliwościach 475 s^{-1} powstała fala stojąca. Odległość jej sąsiednich węzłów równa się 1,5 m. Jaka jest prędkość rozchodzenia się fali w tym ośrodku?

Odp. 1425 m/s.

ROZDZIAŁ 12

Fale głosowe

12.1. Wrażenia słuchowe

Wrażenia słuchowe powstają dzięki falom sprężystym, należącym do pewnego określonego przedziału częstotliwości, dochodzącym od źródła głosu przez sprężysty ośrodek do ucha ludzkiego. W próżni fala głosowa nie może się rozchodzić. Graniczne częstotliwości słyszalne przez ucho ludzkie stanowią tzw. dolną i górną granicę słyszalności i wynoszą odpowiednio około 20 i 20 000 Hz. Te graniczne częstotliwości są w pewnym stopniu zależne od takich czynników, jak muzykalność, wiek itp.

Źródło głosu drgające z częstotliwością „słyszalną", znajdujące się w ośrodku sprężystym stałym, ciekłym lub gazowym, wzbudza w nim falę głosową podłużną. Z mechanizmu powstawania fal podłużnych omawianego w § 11.3 wynika, że falę podłużną można traktować jako falę przemieszczeń cząstek prowadzącą do wytwarzania się zgęszczeń i rozrzedzeń lub też jako falę zmian ciśnień. Jak określiliśmy w § 11.3 między przemieszczeniami i ciśnieniami istnieje przesunięcie w fazie o $\pi/2$. A zatem jeśli w równaniu fali przemieszczeń występuje funkcja sinus, to w równaniu fali ciśnień wystąpi cosinus.

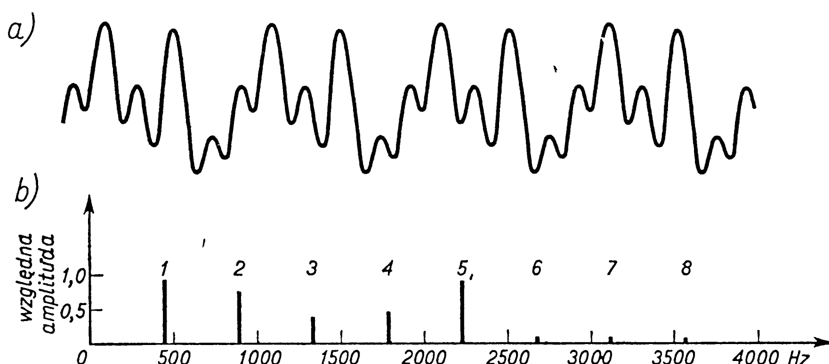
Wśród wrażeń słuchowych rozróżniamy tony, dźwięki i szmery. Tony odpowiadają drganiom harmonicznym źródeł o jednej, ściśle określonej częstotliwości. Dźwięki powstają wtedy, gdy źródło prócz fali podstawowej, o częstotliwości najmniejszej, wysyła fale harmoniczne o częstotliwościach będących całkowitymi wielokrotnościami częstotliwości fali podstawowej. Szmery zaś są to wrażenia słuchowe powstające wtedy, gdy do ucha dochodzą fale o różnych, dowolnych częstotliwościach. Nakładanie się bardzo dużej liczby tonów o dowolnej charakterystyce prowadzi do wrażenia hałasu. Tony proste występują niesłychanie rzadko. Dźwięk wydawany przez kamerton jest zbliżony do tonu prostego, gdyż jego składowe drgania harmoniczne są stosunkowo bardzo słabe w porównaniu z tonem podstawowym.

Dźwięki mogą się różnić wysokością, natężeniem i barwą. Omówimy odpowiedniki fizyczne tych trzech znamion dźwięku.

Wysokość dźwięku jest związana z częstotliwością drgań źródła: częstotliwościom małym odpowiadają dźwięki niskie i odwrotnie. O wysokości decyduje częstotliwość tonu podstawowego.

Natężenie dźwięku mierzy się ilością energii przenoszonej w jednostce czasu przez jednostkę powierzchni ustawionej prostopadle do promienia fali. Natężenie dźwięku wyraża się w $J/m^2 \cdot s$ lub w W/m^2. Ponieważ natężenie dźwięku jest proporcjonalne do ilości energii, a energia ruchu drgającego jest proporcjonalna do kwadratu amplitudy (por. § 10.1), więc tym samym obowiązuje *proporcjonalność natężenia do kwadratu amplitudy drgań źródła.*

Trzecia cecha dźwięku, barwa dźwięku, wiąże się z jego złożonością. O *barwie* odbieranego *dźwięku* decydują: liczba składowych tonów harmonicznych i stosunki ich natężeń. Różnice barwy można ustalać za pomocą słuchu, a można też badać je obiektywnie porównując krzywe zależności wychyleń źródła od położenia równowagi od czasu. Taka krzywa w przypadku źródła tonu prostego ma kształt sinusoidy. Dźwiękom o różnych barwach odpowiadają krzywe o różnych kształtach, o przebiegu okresowym, ale nie sinusoidalnym. Zgodnie z *twierdzeniem Fouriera* (por. § 11.12) takie drgania złożone można rozkładać na szereg drgań harmonicznych. Ustalając liczbę występujących w danym dźwięku tonów harmonicznych i stosunki ich natężeń otrzymujemy charakterystykę danego dźwięku w postaci tzw. *widma akustycznego.*



Rys. 12.1

Na rysunku 12.1a widzimy okresową krzywą drgań struny skrzypcowej wydającej dźwięk, w którym tonem podstawowym jest ton *A* o częstotliwości 440 Hz. Rys. 12.1b przedstawia widmo tego dźwięku. Na osi odciętych odłożone są częstotliwości, a na osi rzędnych — względne amplitudy. Amplituda tonu podstawowego przyjęta jest jako jednostka. Z wykresu widać, że w rozważanym dźwięku występuje osiem tonów harmonicznych, przy czym piąty harmoniczny o częstotliwości 2200 Hz odznacza się dużą amplitudą, a więc i dużym natężeniem.

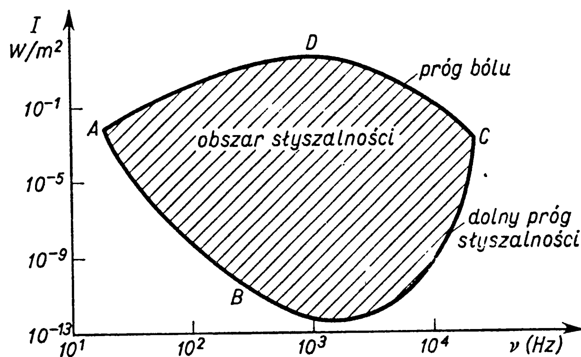
Może warto podkreślić, że krzywe drgań akustycznych można uzyskać za pomocą mikrofonu dołączonego do oscylografu. Nowoczesne przyrządy elektryczne — analizatory dźwięku — rejestrują też widma dźwięku. Zasada ich działania polega na zamianie sygnału dźwiękowego na prąd mikrofonowy skierowany po wzmocnieniu do urządzenia złożonego z całego szeregu filtrów, z których każdy przystosowany jest do przepuszczania drgania o określonej częstotliwości. Drgania przepuszczone przez

odpowiedni filtr wywołują na ekranie oscylografu odchylenia plamki świetlnej proporcjonalne do swojej amplitudy. Dźwięki, których krzywe drgań mają charakter okresowy, a więc mogą być rozłożone na drgania harmoniczne, mają widma liniowe (rys. 12.1b), natomiast wszelkie hałasy, szmery itp. mają widma ciągłe.

W niniejszym punkcie zajmowaliśmy się szerzej falami podłużnymi mechanicznymi z zakresu częstotliwości słyszalnych. W przyrodzie występują ponadto fale podłużne mechaniczne o częstotliwościach mniejszych od 20 Hz, zaliczane do kategorii tzw. *infradźwięków*, oraz fale o częstotliwościach większych od 20 000 Hz, tzw. *ultradźwięki*. Do infradźwięków można zaliczyć np. omawianą w § 11.7 falę podłużną, powstającą w czasie trzęsienia Ziemi. Ultradźwiękom, ich dokładniejszej charakterystyce, metodom ich powstawania oraz zastosowaniom poświęcony jest § 12.7.

12.2. Krzywa czułości ucha ludzkiego

Ucho ludzkie nie jest jednakowo wrażliwe na wszystkie częstotliwości fal głosowych w całym zakresie słyszalności. Badania wykazały, że największa czułość ucha przypada na częstotliwości paru tysięcy herców. Widać to wyraźnie z przebiegu tzw. krzywej czułości ucha ludzkiego (rys. 12.2).



Rys. 12.2

Na osi rzędnych odłożone są natężenia w W/m², a na osi odciętych — częstotliwości ν w Hz. Rzędne krzywej *ABC* przedstawiają najmniejsze natężenia, przy których dźwięk jest jeszcze słyszalny i wyznaczają tzw. *dolny próg słyszalności* lub krócej — *próg słyszalności*. Najmniejsze wartości progu słyszalności odpowiadają częstotliwościom około 3 kHz. Rzędne krzywej *ADC* przedstawiają największe natężenia, przy których dźwięk jest jeszcze słyszalny, ale już z jego odbiorem wiąże się wrażenie bólu. Te maksymalne, graniczne natężenia wyznaczają tzw. *górny próg słyszalności* lub *próg bólu*. Po przekroczeniu tych granicznych natężeń następuje uszkodzenie organu słuchu. Największe wartości progu bólu odpowiadają częstotliwości około 1 kHz. Cały obszar objęty krzywą czułości *ABCDA* nazywamy *obszarem słyszalności*.

Analizując krzywą czułości ucha można stwierdzić dużą rozpiętość natężeń głosu w odniesieniu do tych częstotliwości, na które ucho ludzkie jest najbardziej wrażliwe.

Natężenie tonu o częstotliwości rzędu 1 kHz wzrasta około 10^{13} razy przy przejściu od progu słyszalności do progu bólu. Ucho ludzkie jest więc w tej dziedzinie przyrządem odbiorczym o niesłychanie szerokim zakresie.

Wobec tak szerokich granic natężeń głosu, które ucho ludzkie zdolne jest odbierać, wygodnie było przyjąć pewne natężenie głosu jako natężenie odniesienia i wszystkie inne z nim porównywać stosując skalę logarytmiczną. Jako natężenie odniesienia I_0 wybrano wartość zbliżoną do progu słyszalności tonu o częstotliwości 1 kHz. Wartość I_0 równa się 10^{-12} W/m^2.

Nową wielkość L, charakteryzującą w skali logarytmicznej natężenie I danego dźwięku w stosunku do natężenia I_0 tonu odniesienia, definiuje równanie

$$L = 10 \lg \frac{I}{I_0}. \tag{12.1}$$

Wielkość L nazwano poziomem ciśnienia akustycznego (poziomem natężenia głosu).

Uwzględniając zależność (11.24) otrzymujemy

$$\lg \frac{I}{I_0} = \lg \frac{(p_0')^2}{(p_0^0)^2} = 2 \lg \frac{(p_0')}{(p_0^0)},$$

gdzie p_0' oznacza amplitudę ciśnienia akustycznego badanego dźwięku, p_0^0 — amplitudę ciśnienia akustycznego tonu odniesienia o częstotliwości 1 kHz równą $2 \cdot 10^{-5}$ N/m^2. A zatem wzór (12.1) można przepisać w postaci

$$L = 10 \lg \frac{I}{I_0} = 20 \lg \frac{p_0'}{p_0^0}. \tag{12.2}$$

Jednostką poziomu ciśnienia akustycznego jest 1 decybel (dB).

$$1\,\mathrm{dB} = 20 \lg \frac{p_0'}{2 \cdot 10^{-5}\mathrm{N/m^2}}.$$

Zmiana poziomu ciśnienia akustycznego o jeden decybel wiąże się z najmniejszą uchwytną dla ucha ludzkiego zmianą natężenia głosu.

Różnicę poziomów ciśnienia akustycznego dwóch odbieranych dźwięków można przedstawić jako

$$L_1 - L_2 = 10\left(\lg \frac{I_1}{I_0} - \lg \frac{I_2}{I_0}\right) = 10 \lg \frac{I_1}{I_2} = 20 \lg \frac{p_0'}{p_0''}.$$

Jeśli więc np. różnica poziomów ciśnienia akustycznego wynosi 40 dB, znaczy to, że drugi dźwięk ma natężenie 10^4 razy mniejsze, a ciśnienie akustyczne 10^2 razy mniejsze niż dźwięk pierwszy. Kilka przykładowych wartości L podano w tab. 12.1.

Ucho ludzkie umożliwia porównywanie tzw. subiektywnej głośności tonów o różnej częstotliwości. Wrażenie jednakowej subiektywnej głośności przy odbiorze tonów o różnej częstotliwości występuje na ogół przy niejednakowych poziomach ciśnienia akustycznego. Tak np. ton o częstotliwości $\nu = 50$ Hz i $L = 78$ dB jest równie głośny jak ton o $\nu = 1$ kHz i $L = 60$ dB.

Tabela 12.1

Poziomy ciśnienia akustycznego różnych dźwięków w decybelach

szept ledwo słyszalny	0
szelest liści	10–20
zwykła rozmowa	około 40
głośna rozmowa	60–70
hałas ulicy wielkiego miasta	80–90
hałas hali fabrycznej	90–110
hałas w kabinie samolotu pasażerskiego	około 90
fortissimo orkiestry	90–100

Jako ton odniesienia przy wyrównywaniu głośności przyjęty jest — podobnie jak poprzednio — ton o częstotliwości 1 kHz. Poziom ciśnienia akustycznego tonu odniesienia po doprowadzeniu do wyrównania jego subiektywnej głośności z subiektywną głośnością tonu badanego nosi nazwę poziomu głośności. Aczkolwiek w zasadzie ten poziom ciśnienia jest wyrażony w decybelach, to jednak dla podkreślenia, że tym razem podawana wartość charakteryzuje nie poziom ciśnienia akustycznego, lecz inną wielkość, a mianowicie *poziom głośności*, podaje się ją w *fonach*. A zatem — nawiązując do podanego przykładu — poziom głośności tonu o $\nu = 50$ Hz i $L = 78$ dB wynosi 60 fonów.

12.3. Prędkość fal głosowych

Fale głosowe rozchodzą się w różnych ośrodkach z rozmaitymi prędkościami. Odpowiednie wartości dla kilku ośrodków stałych, ciekłych i gazowych podane są są w tab. 12.2.

Tabela 12.2

Prędkości rozchodzenia się głosu w m/s

1. Gazy w warunkach normalnych	
powietrze	331,3
wodór	1286
tlen	317,2
dwutlenek węgla	261
2. Ciecze w temperaturze 17°C	
woda	1430
rtęć	1460
alkohol etylowy	1440
3. Ciała stałe w temperaturze 20°C	
miedź	3560
szkło	4000–5600
żelazo	5130
aluminium	5100

Do obliczenia prędkości rozchodzenia się głosu można wykorzystywać wzór Newtona:

$$v = \sqrt{\frac{\text{moduł sprężystości}}{\text{gęstość}}},$$

podstawiając odpowiedni moduł sprężystości ośrodka.

Przypominamy, że zjawisko rozchodzenia się fal głosowych w ośrodkach gazowych traktujemy jako zjawisko adiabatyczne. A zatem prędkość rozchodzenia się głosu w gazach obliczamy na podstawie wzoru

$$v = \sqrt{\frac{\varkappa p}{\varrho}},$$

gdzie $\varkappa = c_p/c_V$, p oznacza ciśnienie gazu, ϱ — jego gęstość.

Zadanie. Znając zależność między długością fali λ, prędkością v i częstotliwością ν obliczyć graniczne długości fal głosowych, odpowiadające dolnej i górnej granicy słyszalności w powietrzu w temperaturze pokojowej.

Podstawowemu tonowi *A* o częstotliwości $\nu = 440$ Hz w przybliżeniu odpowiada długość fali $\lambda = 0{,}80$ m w powietrzu, w temperaturze pokojowej pod normalnym ciśnieniem.

12.4. Podstawowe zjawiska towarzyszące rozchodzeniu się fal głosowych

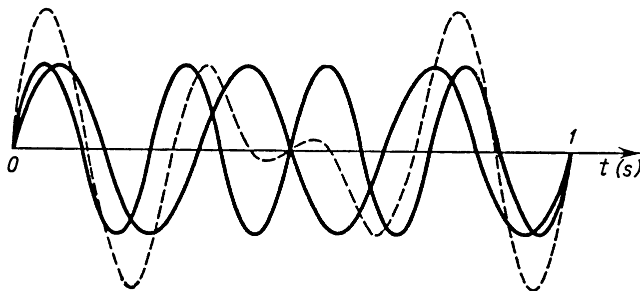
1. *Odbicie fal głosowych.* Gdy fala głosowa napotyka na swej drodze nieprzenikliwą przeszkodę, ulega odbiciu według ogólnych praw odbicia fal (por. § 11.4). Czasem zjawisku odbicia głosu towarzyszy wytwarzanie się echa, gdy fala głosowa po odbiciu powraca do ucha obserwatora. Z odstępu czasu, upływającego między nadaniem fali głosowej i jej odbiorem w postaci echa, można wnioskować o odległości przeszkody odbijającej od ucha obserwatora.

2. *Załamanie fal głosowych.* Przejściu fali głosowej do ośrodka o innej prędkości rozchodzenia się dźwięku towarzyszy na ogół zjawisko załamania, podlegające również ogólnym prawom załamania (§ 11.4).

3. *Interferencja fal głosowych.* Jeżeli w danym ośrodku spotykają się dwie lub więcej fal głosowych o jednakowej częstotliwości, występuje zjawisko interferencji. W wyniku interferencji można otrzymać wzrost lub zmniejszenie wychyleń poszczególnych cząstek ośrodka od położenia równowagi, zależnie od różnicy faz nakładających się fal. Zwiększenie amplitudy drgań odpowiada wzrostowi natężenia tonu. Zjawisko interferencji fal omawialiśmy ogólnie w § 11.8. Szczególnym przypadkiem interferencji fal głosowych jest wytwarzanie się fal stojących, którym poświęcony jest § 12.5.

4. *Dudnienia.* Podczas nakładania się fal dźwiękowych o nieznacznie różniących się częstotliwościach otrzymujemy charakterystyczne zjawisko znane w akustyce pod nazwą *dudnień*. Na rys. 12.3 przedstawione jest nakładanie się dwóch fal o jednakowych amplitudach i o częstotliwościach odpowiednio równych 3 i 4 Hz. Do sporządzenia wykresu wybrane zostały częstotliwości nieakustyczne dla uzyskania przejrzystości ujęcia graficznego. Wnioski można uogólnić na fale o częstotliwościach akustycznych. Linia kreskowana na rysunku przedstawia wynik nakładania się fal w ciągu jednej sekundy. Fala wypadkowa (linia kreskowana) miałaby w następnych sekundach przebieg ten sam,

gdyż zjawisko periodycznie się powtarza. Z rysunku widać, że w każdej sekundzie mamy jednorazowy utrzymujący się przez pewien czas wzrost amplitudy i jednokrotny, również utrzymujący się przez pewien czas spadek amplitudy prawie do zera. Częstotliwość tych wzmocnień i osłabień odpowiada różnicy częstotliwości obu nakładających się fal. Gdybyśmy zrobili wykres dwóch nakładających się fal akustycznych, np. o częstotliwościach 435 Hz i 440 Hz, to takich wzrostów i osłabień otrzymalibyśmy pięć na sekundę. Wzrost amplitudy drgań akustycznych odpowiada wzrostowi natężenia odbieranego dźwięku, zmniejszenie amplitudy wiąże się z osłabieniem dźwięku. Okresowo powtarzające się przejścia od wzrostów natężenia dźwięku do zaniku tworzą właśnie zjawisko dudnień. Liczba dudnień przypadających na sekundę, czyli ich częstotliwość, odpowiada różnicy częstotliwości nakładających się fal.



Rys. 12.3

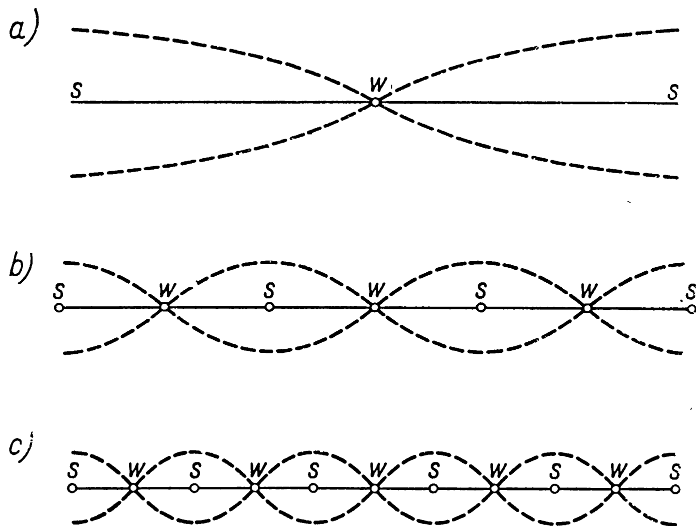
Efekt dudnień można zaobserwować np. przy równoczesnym uderzeniu dwóch sąsiednich klawiszów pianina. Przeciętne ucho ludzkie reaguje na dudnienia aż do różnicy częstotliwości obu źródeł wynoszącej 7 Hz.

Doświadczalnie można wykazać istnienie dudnień za pomocą dwóch kamertonów o jednakowej częstotliwości. Nakładając metalową nasadkę na ramię jednego z kamertonów zmieniamy nieco jego częstotliwość. Dzięki temu po pobudzeniu do drgań obu kamertonów słyszymy dudnienia. Obliczając liczbę dudnień na sekundę można ustalić wywołaną zmianę częstotliwości kamertonu. Metodę dudnień wykorzystuje się niekiedy do wyznaczania częstotliwości drgań dowolnych źródeł (nieakustycznych) dobierając do nich taki układ drgający o znanej częstotliwości, aby wystąpiły dudnienia o częstotliwości małej, łatwej do policzenia.

12.5. Fale głosowe stojące

Jak pamiętamy, szczególnym przypadkiem interferencji jest powstawanie fal stojących. Musi być przy tym zachowany warunek równych częstotliwości i zbliżonych amplitud nakładających się fal, rozchodzących się w przeciwnych kierunkach. Najłatwiej jest osiągnąć te warunki podczas interferencji fali padającej i fali odbitej. W zależności od charakteru ośrodka odbijającego w punkcie odbicia wytworzyć się musi węzeł lub strzałka. Węzeł fali stojącej tworzy się przy odbiciu od ośrodka bardziej bezwładnego, bardziej masywnego, strzałka zaś na granicy ośrodka mniej bezwładnego.

1. Jako pierwszy przykład wytwarzania się fal głosowych stojących rozważmy falę powstającą w pręcie zamocowanym poziomo, dokładnie w połowie długości. Przez pocieranie podłużne tego pręta wprawiamy go w drgania, którym towarzyszy wysoki ton. Zaburzenie wywołane w pręcie dochodząc do drugiego końca ulega odbiciu od otaczającego powietrza i wraca interferując z ciągiem fal padających, rozchodzących się od pobudzanego końca pręta. W wyniku interferencji ciągów fal padających i odbitych wytwarza się fala stojąca, której strzałki wypadają na końcach pręta, a węzeł w miejscu zamocowania. Cała długość pręta stanowi więc pół długości fali. Obraz tej fali stojącej w *dwóch chwilach różniących się o pół okresu* przedstawiają linie kreskowane na rys. 12.4. Należy jednak pamiętać, że ten rysunek nie jest bezpośrednim odzwierciedleniem drgań zachodzących w pręcie, gdyż w nim rozchodzi się fala podłużna. Wykonując rys. 12.4 wykorzystujemy to samo założenie co przy rys. 11.7, a mianowicie wychylenia podłużne cząstek pręta odkładamy w kierunku poprzecznym. W ten sposób widać, że punktowi W (węzłowi) odpowiada wychylenie zerowe, a punktom S (strzałkom) odpowiadają wychylenia maksymalne.



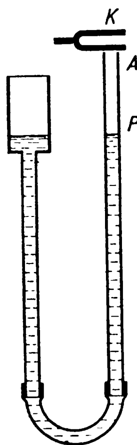
Rys. 12.4

Prócz fali stojącej zaznaczonej na rys. 12.4a możliwe jest wytworzenie się fali stojącej o większej liczbie węzłów i strzałek, jak np. na rys. 12.4b i 12.4c, z tym jednak, że w miejscu zamocowania, czyli w środku pręta, stale utrzymuje się węzeł, a na końcach strzałki.

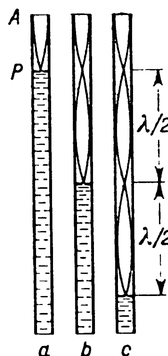
W odniesieniu do fal przedstawionych na rys. 12.4a, b, c długość pręta stanowi odpowiednio $\frac{1}{2}\lambda$, $\frac{3}{2}\lambda$ i $\frac{5}{2}\lambda$. Długości poszczególnych fal maleją w stosunku 1:3:5. Stąd wniosek, że częstotliwości odpowiednich fal mają się do siebie jak 1:3:5, przy czym najmniejsza jest częstotliwość tonu odpowiadającego rys. 12.4a. Jest to najniższy, a równocześnie najsilniejszy ton wydawany przez pocierany pręt. Jest on właśnie to-

nem podstawowym. Tony pozostałe, tj. tony wyższe harmoniczne są znacznie słabsze. W rozważanym przypadku powstają tylko drgania harmoniczne nieparzyste. Częstotliwość tonu podstawowego związana z powstawaniem w danym ciele fali stojącej o największej długości stanowi tzw. *częstotliwość drgań własnych układu drgającego.*

2. Drugi przykład wytwarzania się fali głosowej stojącej jest przedstawiony na rys. 12.5. Drgający kamerton K umieszczamy u wylotu A długiej rury, stanowiącej ruchome ramię naczyń połączonych. Przesuwając to ramię w górę lub w dół, zmieniamy położenie powierzchni swobodnej P cieczy względem wylotu rury. W rurze następuje interferencja fali padającej, rozchodzącej się od kamertonu przez słup powietrza do menisku cieczy i fali odbitej, powracającej od menisku cieczy do wylotu rury i tam ulegającej ponownemu odbiciu od otaczającego powietrza. Odbicie od menisku cieczy — to odbicie od ośrodka bardziej bezwładnego. Odbicie od otaczającego powietrza traktujemy jako odbicie od ośrodka mniej bezwładnego. Jeśli więc tak dobierzemy położenie ruchomego ramienia, by odległość od wylotu rury do menisku cieczy była równa nieparzystej wielokrotności ćwiartek długości fali, to w pobliżu wylotu trwale utrzymywać się będzie strzałka drgań, a na powierzchni swobodnej cieczy — węzeł, a więc powstanie fala stojąca. Długość słupa powietrza nad cieczą powinna się równać $\frac{1}{4}\lambda$, $\frac{3}{4}\lambda$, $\frac{5}{4}\lambda$ itd. (rys. 12.6a, b, c).



Rys. 12.5



Rys. 12.6

Wytworzenie się fali stojącej w rurze ściśle się wiąże ze zjawiskiem rezonansu: drgania w słupie powietrza w rurze zostają wymuszone przez drgania kamertonu. Wobec małego oporu ośrodka można uważać, że częstotliwość wymuszająca i częstotliwość rezonansowa są sobie równe. A zatem długość fali λ w powietrzu jest związana z częstotliwością drgań kamertonu i prędkością głosu v w powietrzu w danej temperaturze zależnością

$$\lambda \nu = v.$$

Doświadczalnie stwierdzamy osiągnięcie odpowiedniego położenia menisku cieczy słysząc znaczne wzmożenie dźwięku. Tym razem odbieramy równocześnie dźwięk po-

chodzący od kamertonu i od drgań własnych rezonującego słupa powietrza zawartego w rurze. W związku z tym, że u wylotu rury występują pewne dodatkowe zaburzenia, powodujące przesunięcie położenia strzałki fali stojącej, pomiar długości fali przeprowadzamy zazwyczaj w ten sposób, że w czasie przesuwania się menisku cieczy w rurze notujemy dwa jego położenia, odpowiadające kolejnym wzmocnieniom odbieranego dźwięku. Odległość tych położeń stanowi połowę długości fali. Znając długość fali i częstotliwość drgań kamertonu można obliczyć prędkość rozchodzenia się głosu w powietrzu w danej temperaturze. Wypełniając rurę innym gazem wyznaczamy podobnie prędkość rozchodzenia się głosu w innym ośrodku gazowym w temperaturze doświadczenia.

3. Z falami stojącymi mamy także do czynienia w przypadku drgań strun. Długość fali tonu podstawowego struny o długości L zamocowanej na obu końcach odpowiada podwójnej długości struny

$$\lambda = 2L.$$

Wtedy na strunie wytwarza się fala stojąca, mająca dwa węzły na końcach, a strzałkę w środku. Łatwo sprawdzić, że tonom wyższym harmonicznym odpowiadają długości fal

$$\lambda_n = \frac{2L}{n}, \quad \text{gdzie} \quad n = 1, 2, 3, \ldots$$

Prędkość rozchodzenia się fali w strunie (wyrażona w układzie SI) może być przedstawiona następującym wzorem:

$$v = \sqrt{\frac{F}{S\varrho}},$$

gdzie F oznacza siłę rozciągającą strunę (w N), S — pole przekroju poprzecznego struny (w m^2), ϱ — gęstość struny (w kg/m^3).

Zwracamy uwagę, że struna o danej długości L może mieć różne częstótliwości własne, może więc wydawać różne tony podstawowe w zależności od zastosowanego napięcia struny (strojenie strun). Częstotliwość drgań harmonicznych struny wyraża się wzorem

$$\nu = \frac{v}{\lambda} = \frac{n}{2L}\sqrt{\frac{F}{S\varrho}}.$$

4. W piszczałkach różnych typów także istotną rolę odgrywają fale stojące. Rozróżniamy piszczałki otwarte i zamknięte. Obraz drgań podłużnych słupa powietrza w piszczałce (jednostronnie) zamkniętej przedstawia rys. 12.7.

Rysunek 12.7a odpowiada drganiom podstawowym piszczałki zamkniętej. Częstotliwość tych drgań wyraża się wzorem

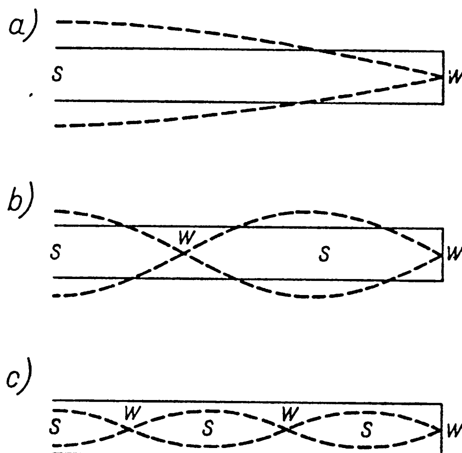
$$\nu = \frac{v}{4L},$$

gdzie v oznacza prędkość głosu w powietrzu, L — długość piszczałki.

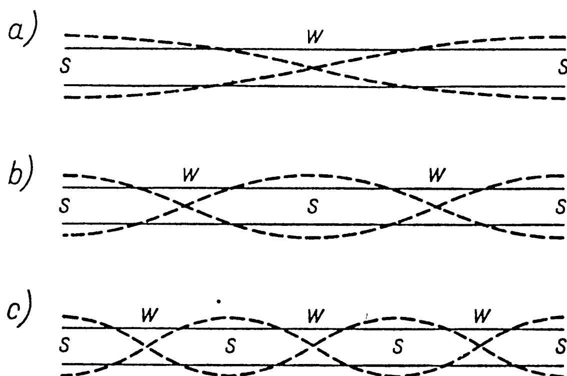
Rysunki 12.7b i c przedstawiają drgania harmoniczne o częstotliwościach 3 i 5 razy większych od częstotliwości drgania podstawowego. Piszczałka zamknięta wytwarza jedynie drgania harmoniczne nieparzyste.

Powstanie drgań słupa powietrza w piszczałcę obustronnie otwartej przedstawia schematycznie rys. 12.8. Częstotliwość tonu podstawowego tym razem wyraża się wzorem

$$\nu = \frac{v}{2L}.$$



Rys. 12.7

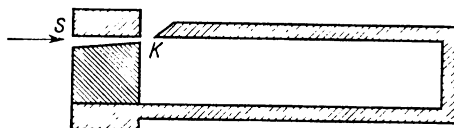


Rys. 12.8

Rysunki 12.8b i c przedstawiają tworzenie się tonów harmonicznych wyższych. W piszczałce otwartej występują zarówno parzyste, jak i nieparzyste tony harmoniczne.

Na rysunku 12.9 wyjaśniony jest sposób wywołania drgań w piszczałkach. Prąd powietrza wdmuchiwany jest do piszczałki przez wąski kanał S. Jest on kierowany na ostry, klinowaty brzeg K otworu wyciętego w ścianie bocznej piszczałki. Rozbiciu

strumienia powietrza na krawędzi K towarzyszy wystąpienie syczenia. Syczenie jest mieszaniną rozmaitych tonów. Słup powietrza zawarty w piszczałce odpowiada (rezonuje) na te tony, które mają częstotliwości zgodne z jego częstotliwościami własnymi. Dzięki temu piszczałka staje się źródłem silnych drgań akustycznych. Ten koniec piszczałki, przez który wdmuchiwane jest powietrze, zachowuje się jak otwarty koniec piszczałki.



Rys. 12.9



Rys. 12.10

5. W akustyce powszechnie znane jest doświadczenie Kundta z falami stojącymi. Poziomo umieszczony pręt P (rys. 12.10) zamocowany jest dokładnie w środku. Pocierając podłużnie jeden koniec pręta wzbudzamy w nim falę stojącą. Drugi koniec pręta zakończony tłoczkiem K wprowadzony jest do wnętrza szklanej rury, której dno posypane jest cienką warstwą lekkiego proszku, np. tartego korka lub lycopodium (zarodniki widłaku). Drgający tłok K jest źródłem drgań dla zawartego w rurze słupa powietrza. Fala podłużna przechodzi przez powietrze wewnątrz rury i ulega odbiciu od tłoczka T zamykającego drugi wylot. Dzięki temu w rurze powstają ciągi fal biegnących i odbitych, które interferują ze sobą. Nie zawsze jednak w rurze wytwarza się fala stojąca. Odpowiednie warunki do jej powstania uzyskujemy przesuwając powoli tłoczek T w czasie pobudzania pręta P do drgań. Przy pewnym położeniu tłoczka można zaobserwować zaburzenie w układzie proszku na dnie rury. Wyraźnie odcinają się od siebie takie miejsca, gdzie proszek pozostaje w spoczynku oraz takie, gdzie cząstki proszku, a więc i cząstki powietrza w rurze, podlegają silnym drganiom. Są to węzły i strzałki powstającej w rurze fali stojącej. W pobliżu powierzchni tłoczków zamykających rurę powstają węzły ruchu cząstek, a jednocześnie strzałki ciśnienia, jak to wynika z teorii fal stojących (§ 11.10). Bez trudu można wymierzyć odległości między węzłami i obliczyć długość fali stojącej, powstającej w powietrzu. Warto podkreślić, że źródłem tej fali były drgania wzbudzone w pręcie, a zatem częstotliwość drgań słupa powietrza równa się częstotliwości drgań podstawowych pręta. Zachodzi tu bowiem znowu *zjawisko rezonansu.*

Korzystając z zależności $\lambda = v/\nu$ możemy napisać

$$\lambda_{\text{w pręcie}} = \frac{v_{\text{w pręcie}}}{\nu_{\text{w pręcie}}},$$

$$\lambda_{\text{w powietrzu}} = \frac{v_{\text{w powietrzu}}}{\nu_{\text{w powietrzu}}}.$$

Wobec tego, że częstotliwość drgań w pręcie i w powietrzu jest taka sama, otrzymujemy

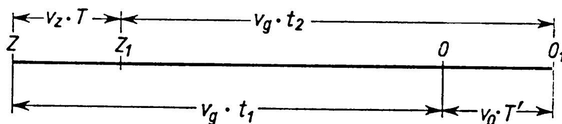
$$\frac{\lambda_{\text{w pręcie}}}{\lambda_{\text{w powietrzu}}} = \frac{v_{\text{w pręcie}}}{v_{\text{w powietrzu}}}.$$

Doświadczalnie łatwo znajdujemy długość fali w pręcie i długość fali w powietrzu. A zatem, znając prędkość rozchodzenia się fali podłużnej w powietrzu, można z tego równania obliczyć prędkość rozchodzenia się fali w materiale pręta.

12.6. Zjawisko Dopplera

Ani zjawisko odbicia, ani zjawisko załamania nie mają wpływu na częstotliwość odbieranego tonu. Zmiana częstotliwości tonu odbieranego w stosunku do częstotliwości tonu nadawanego występuje w przypadku ruchu źródła, ruchu obserwatora lub też obu na raz względem ośrodka, w którym rozchodzi się fala głosowa. Mimo że zjawisko to, znane pod nazwą zjawiska Dopplera, występuje we wszystkich rodzajach ruchu falowego, omówimy je dokładnie w odniesieniu do fal głosowych, gdyż w tej dziedzinie znamy je ogólnie z naszych codziennych obserwacji.

Każdy z nas niejednokrotnie miał okazję stwierdzić dość gwałtowną zmianę wysokości odbieranego tonu w momencie, gdy poruszające się źródło dźwięku nas wymija. Zmienia się wysokość tonu gwizdka przejeżdżającego obok nas parowozu, zmienia się wysokość tonu syreny mijającego nas samochodu itp.



Rys. 12.11

Następujące rozumowanie pozwoli ustalić zależność między częstotliwością ν tonu wysyłanego, częstotliwością ν' tonu odbieranego oraz prędkościami: źródła (v_z), obserwatora (v_0) i głosu (v_g).

Załóżmy początkowo, że źródło głosu (rys. 12.11) i obserwator poruszają się wzdłuż tej samej prostej w tym samym kierunku. Niech punkt Z przedstawia położenie źródła w chwili nadania pierwszego sygnału akustycznego. Punkt O niech wyobraża położenie obserwatora w chwili odbioru tego sygnału. Odległość ZO jest drogą przebytą przez falę głosową w pewnym czasie t_1, a zatem

$$ZO = v_g t_1.$$

Drugi z kolei sygnał akustyczny jest wysłany przez źródło po czasie T (okres fal na-

dawanych) z nowego położenia Z_1. Odległość ZZ_1 jest drogą przebytą przez źródło w czasie T, a więc

$$ZZ_1 = v_z T.$$

Ten drugi sygnał dogania obserwatora po czasie t_2 w nowym położeniu O_1. Odległość $Z_1 O_1$ równa się więc

$$Z_1 O_1 = v_g t_2.$$

Czas upływający między odbiorem pierwszego i drugiego sygnału jest okresem drgań odbieranych, który oznaczamy przez T'. W tym czasie obserwator przesunął się o odległość OO_1, przy czym

$$OO_1 = v_0 T'.$$

Spoglądając na rysunek łatwo sprawdzić, że odległość ZO_1 można wyrazić dwojako: jako sumę $ZZ_1 + Z_1 O_1$ lub jako sumę $ZO + OO_1$. Innymi słowy,

$$v_z T + v_g t_2 = v_g t_1 + v_0 T',$$

skąd

$$v_g(t_2 - t_1) = v_0 T' - v_z T. \tag{12.3}$$

Czas upływający między chwilą nadania pierwszego sygnału i chwilą odbioru drugiego sygnału również możemy wyrazić dwojako: jako sumę $t_1 + T'$ lub jako sumę $T + t_2$, a zatem

$$t_1 + T' = T + t_2,$$

skąd

$$t_2 - t_1 = T' - T.$$

Wstawiając to wyrażenie do równania (12.3) otrzymujemy

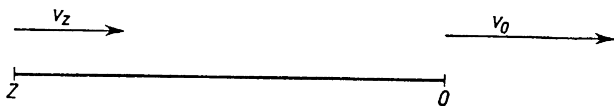
$$v_g(T' - T) = v_0 T' - v_z T,$$

skąd okres drgań odbieranych

$$T' = T \frac{v_g - v_z}{v_g - v_0}.$$

Pamiętając, że okres jest odwrotnością częstotliwości drgań, znajdujemy

$$\nu' = \nu \frac{v_g - v_0}{v_g - v_z}. \tag{12.4}$$



Rys. 12.12

Jest to wzór określający częstotliwość drgań odbieranych w zależności od częstotliwości drgań nadawanych, prędkości rozchodzenia się głosu, prędkości źródła i prędkości obserwatora.

Aby dokładniej zbadać, jak się zmienia częstotliwość drgań odbieranych, rozważmy kolejno następujące przypadki:

1. Niech źródło i obserwator poruszają się w kierunku zgodnym z kierunkiem prędkości głosu od źródła do obserwatora (rys. 12.12). Kierunek prędkości głosu od źródła do obserwatora umówmy się traktować jako dodatni. Jest to właśnie ten przypadek, który rozważaliśmy przy wyprowadzaniu wzoru. A zatem obowiązuje zależność

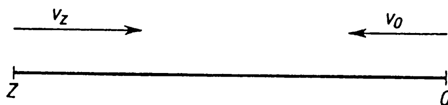
$$\nu' = \nu \frac{v_g - v_0}{v_g - v_z}.$$

Należy jednak rozróżniać trzy możliwości:

a) $v_0 > v_z$, czyli odległość między źródłem a obserwatorem z biegiem czasu rośnie; tym razem licznik $v_g - v_0$ jest zawsze mniejszy od mianownika $v_g - v_z$. W tych warunkach częstotliwość drgań odbieranych jest mniejsza od częstotliwości drgań nadawanych.

b) $v_0 < v_z$, czyli odległość między źródłem i obserwatorem maleje. Z porównania wielkości licznika i mianownika wynika, że tym razem częstotliwość drgań odbieranych jest większa od częstotliwości drgań nadawanych.

c) $v_0 = v_z$. Częstotliwość drgań odbieranych i nadawanych jest jednakowa.



Rys. 12.13



Rys. 12.14

2. Prędkość źródła jest zgodna z kierunkiem prędkości głosu, a więc dodatnia, natomiast prędkość obserwatora skierowana jest przeciwnie, a więc jest ujemna (rys. 12.13). Wzór 12.4 przyjmuje postać

$$\nu' = \nu \frac{v_g + v_0}{v_g - v_z}.$$

W tych warunkach odległość między źródłem i obserwatorem maleje, częstotliwość drgań odbieranych jest większa od częstotliwości drgań nadawanych.

3. Niech v_z będzie mniejsze od zera, a v_0 większe od zera (rys. 12.14). Odległość między źródłem i obserwatorem rośnie. Wzór (12.4) przyjmuje postać

$$\nu' = \nu \frac{v_g - v_0}{v_g + v_z}.$$

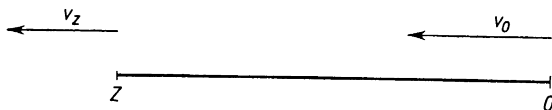
W tym przypadku, jak widać, ν' jest zawsze mniejsze od ν.

4. Tym razem niech obie prędkości v_z i v_0 będą skierowane przeciwnie względem

prędkości głosu v_g, a zatem obie niech będą ujemne (rys. 12.15). Wzór (12.4) przekształca się na

$$\nu' = \nu \frac{v_g + v_0}{v_g + v_z}.$$

Podobnie jak w przypadku pierwszym rozróżniamy trzy możliwości:



Rys. 12.15

a) $v_0 > v_z$ — odległość między źródłem i obserwatorem maleje

$$\nu' > \nu;$$

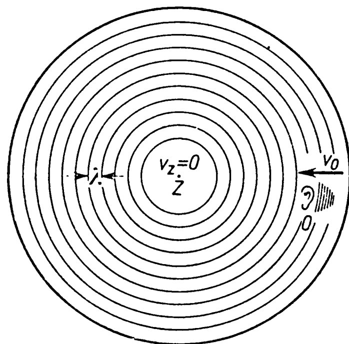
b) $v_0 < v_z$ — odległość między źródłem i obserwatorem rośnie

$$\nu' < \nu;$$

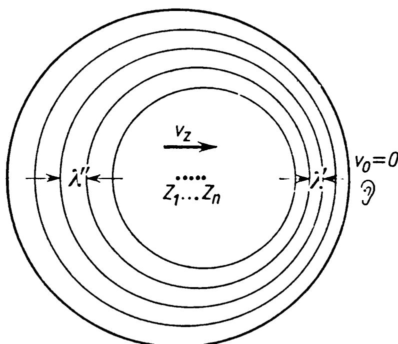
c) $v_0 = v_z$ — odległość między źródłem i obserwatorem nie ulega zmianie

$$\nu' = \nu.$$

Zastosowanie wyprowadzonych wzorów pozwoli Czytelnikowi na ustalenie, czy zmiana częstotliwości odbieranego sygnału będzie jednakowa, gdy: a) źródło zbliża się do nieruchomego obserwatora z prędkością u, b) obserwator zbliża się do nieruchomego źródła z prędkością u?



Rys. 12.16



Rys. 12.17

Odpowiedź na postawione pytania być może jest nieco zaskakująca i wymaga bliższego przeanalizowania. Okazuje się, że zmiany częstotliwości w obu przypadkach są różne. Wiąże się to z tym, że wystąpienie zjawiska Dopplera wynika z ruchu obserwatora i źródła względem ośrodka, a nie z ich ruchu względnego. Jeśli obserwator porusza się np. w stronę nieruchomego źródła (rys. 12.16), to jego ucho odbiera w czasie 1 se-

kundy więcej fal niż wtedy, gdy jest ono nieruchome. Jeżeli jednak obserwator jest nieruchomy, a źródło się do niego zbliża (rys. 12.17), to efekt jest taki, że długość fali dochodzącej do ucha jest mniejsza niż w przypadku źródła nieruchomego. Widać to wyraźnie na rys. 12.17. Źródło w swoim ruchu usiłuje dogonić wysyłane fale, co powoduje zmniejszenie odległości między kolejnymi np. grzbietami, a więc zmniejsza długość fali. Zmniejszeniu długości fali — przy stałej prędkości głosu w ośrodku — odpowiada wzrost częstotliwości. Tak więc inne czynniki wywołują zmianę częstotliwości odbieranego tonu przy ruchu źródła, a inne przy ruchu obserwatora i stąd różne wyniki w obu przypadkach wymienionych w pytaniach.

Kończąc na tym omawianie zjawiska Dopplera raz jeszcze podkreślimy, że chodzi tu o zmianę częstotliwości (wysokości) odbieranego tonu, wynikającą z ruchu obserwatora i źródła. Należy to sobie wyraźnie uświadomić, aby nie kojarzyć błędnie efektu Dopplera ze zmianą natężenia dźwięku, wynikającą ze zmiany odległości obserwatora od źródła. W ośrodku jednorodnym, nie ograniczonym i nie pochłaniającym energii natężenie dźwięku zmienia się odwrotnie proporcjonalnie do kwadratu odległości od źródła fali kulistej (źródło punktowe lub kuliste):

$$\frac{I_1}{I_2} = \frac{r_2^2}{r_1^2}. \qquad (12.5)$$

To prawo nie ma nic wspólnego ze zjawiskiem Dopplera.

P y t a n i e. Jak można uzasadnić wzór (12.5) uwzględniając rozkład energii na wzrastającą powierzchnię falową?

12.7. Ultradźwięki

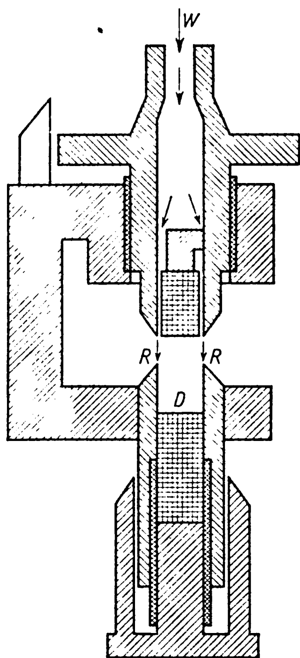
Fale ultraakustyczne o częstotliwościach większych od 20 000 Hz nie wywołują wrażeń słuchowych u człowieka, gdyż, jak pamiętamy, górna granica słyszalności wynosi właśnie około 20 000 drgań na sekundę. Wiemy jednak, że na ultradźwięki reagują niektóre zwierzęta, np. psy i nietoperze. Nietoperz podczas lotu wydaje „ultradźwiękowe piski", które ulegają odbiciu od ciał w otoczeniu, po czym część ich energii wraca w postaci echa ultradźwiękowego do nietoperza. Zależnie od odległości przeszkody echo ultradźwiękowe wraca po dłuższym lub krótszym okresie czasu, co pozwala nietoperzowi na orientowanie się w rozmieszczeniu przeszkód na drodze jego lotu i skuteczne ich omijanie. Wrażliwość psów na sygnały ultradźwiękowe pozwala na kierowanie tresowanych psów takimi niesłyszalnymi dla człowieka sygnałami.

Po tym krótkim wstępie przejdźmy do omówienia sposobów wytwarzania ultradźwięków. Można tu wyróżnić trzy zasadnicze metody: mechaniczną, elektryczną (piezoelektryczną) oraz magnetyczną (magnetostrykcyjną).

Mechaniczne wytwarzanie ultradźwięków sprowadza się do użycia odpowiednio zbudowanych piszczałek i syren. Budowę np. piszczałki Galtona przedstawia rys. 12.18.

W piszczałce Galtona strumień powietrza wdmuchiwany przez wlot *W* pada na ostre brzegi rezonatora *RR*, powodując powstanie drgań o różnych częstotliwościach. Między innymi powstaje silne drganie o częstotliwości odpowiadającej częstotliwości

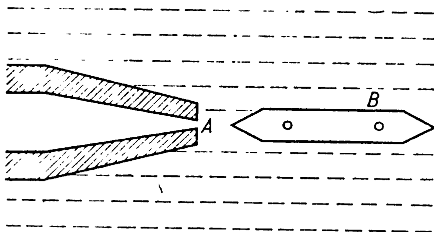
tonu podstawowego rezonatora. Częstotliwość drgań własnych rezonatora można zmieniać przez przesuwanie za pomocą odpowiedniej śruby jego ruchomego dna *D*. W ten sposób można otrzymać intensywne drgania o częstotliwościach ultraakustycznych, dochodzących do 100 tysięcy herców.



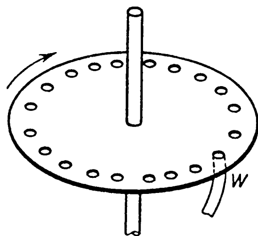
Rys. 12.18

Do praktycznych celów stosuje się zmodyfikowaną piszczałkę Galtona umieszczoną w cieczy (rys. 12.19). Ciecz wypływająca z dużą prędkością z wąskiej dyszy *A* natrafia na stalową blaszkę *B* wprawiając ją w drgania. Przez dobór wielkości i kształtu blaszki oraz odpowiednie umieszczenie jej względem dyszy można uzyskać silne drgania ultradźwiękowe w cieczy. Ten rodzaj źródła dźwięków znalazł zastosowanie w przemyśle przy tworzeniu emulsji w cieczach pod wpływem drgań ultraakustycznych.

Innym rodzajem urządzenia mechanicznego służącego do wytwarzania ultradźwięków są syreny. Zasadę ich działania tłumaczy rys. 12.20. Widzimy tam tarczę osadzoną na osi, mającą szereg otworów. Tarczę wprawiamy w szybki ruch obrotowy, a równocześnie skierowujemy na nią silny strumień powietrza przez wylot *W*. Wskutek obrotu tarczy otrzymujemy rytmiczne zamykanie i otwieranie przepływu powietrza, czemu towarzyszy powstanie fali dźwiękowej lub ultradźwiękowej. Wydawałoby się, że przez zwiększanie liczby otworów w tarczy i powiększanie prędkości obrotu można by w takim urządzeniu w prosty sposób osiągać olbrzymie częstotliwości. Sprawa nie przedstawia się tak prosto m. in. z tego względu, że ze wzrostem prędkości rośnie gwałtownie siła odśrodkowa bezwładności, która może grozić rozerwaniem tarczy. Względy bezpieczeństwa narzucają pewną górną granicę częstotliwości obrotów tarczy. Z uwagi zresztą na dużą absorpcję fal o bardzo dużej częstotliwości w powietrzu na ogół w takich urządzeniach nie dąży się do zbytniego podnoszenia częstotliwości.



Rys. 12.19



Rys. 12.20

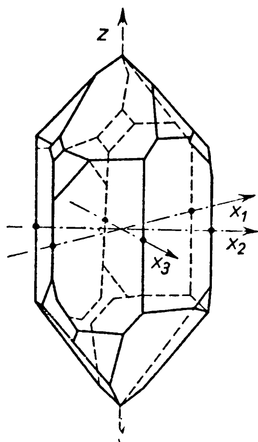
Mechaniczne metody otrzymywania ultradźwięków stosujemy wtedy, gdy nie zależy nam na dużych częstotliwościach, a potrzebne są duże moce przy częstotliwościach średnich. Piszczałka Galtona daje częstotliwości od $3{,}5 \cdot 10^3$ do $100 \cdot 10^3$ Hz, syreny —

kilkadziesiąt tysięcy herców. Ich moc wynosi ok. 2000 W. Może warto porównać tę wartość z mocą akustyczną źródeł dźwięku; tak np. moc akustyczna

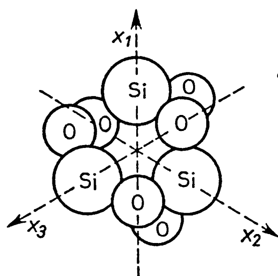
ludzi mówiących szeptem wynosi ok. 10^{-7} W,
ludzi mówiących normalnie — 10^{-5} W,
ludzi krzyczących — 10^{-3} W.

Przejdźmy do drugiego sposobu otrzymywania ultradźwięków. Polega on na wykorzystaniu *zjawiska piezoelektrycznego.*

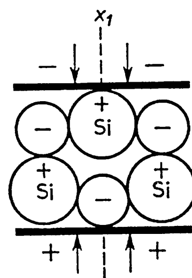
Stwierdzono, że niektóre ciała, m.in. kwarc, sól Seignette'a, turmalin, tytanian baru, mają interesującą właściwość, polegającą na tym, że rozciąganie lub ściskanie płytek, odpowiednio z tych ciał wyciętych, powoduje pojawianie się ładunków elektrycznych na powierzchniach tych płytek.



Rys. 12.21



Rys. 12.22



Rys. 12.23

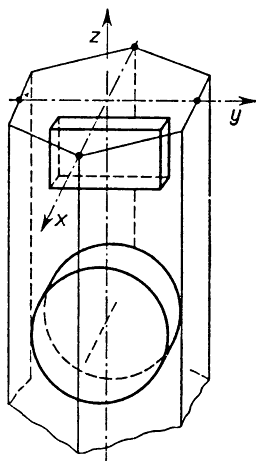
Omówimy to nieco dokładniej na przykładzie kwarcu. Rys. 12.21 przedstawia kryształ kwarcu o budowie jonowej (jony krzemu dodatnie i jony tlenu ujemne). Trwałość budowy krystalicznej wiąże się z oddziaływaniami elektrostatycznymi ładunków. W krysztale takim rozróżniamy oś z potrójnej symetrii, zwaną osią optyczną, oraz osie x drugorzędowe, prostopadłe do osi z, łączące przeciwległe krawędzie boczne. Osie x są to tzw. osie elektryczne.

Rysunek 12.22 przedstawia schematycznie układ jonów Si i O w przestrzeni. W sąsiedztwie każdego jonu Si (o ładunkach dodatnich) znajdują się dwa jony O (o 2 ładunkach ujemnych). W sposób bardziej uproszczony przedstawia tę sytuację rys. 12.23, gdzie każde duże kółko przedstawia jon Si z ładunkiem dodatnim, a każde mniejsze kółko przedstawia dwa jony tlenu z ładunkami ujemnymi. Jak widać, rysunek ostatni przedstawia schemat przekroju prostopadłego do osi optycznej. Wyobraźmy sobie, że taką komórkę krystaliczną objęliśmy dwiema płytkami metalowymi i zastosowaliśmy siłę ściskającą równoległą do osi x_1. Wtedy do górnej płytki zbliżają się wyciśnięte

jak gdyby ze środka ładunki ujemne, do dolnej — ładunki dodatnie. Zastosowanie siły rozciągającej o kierunku równoległym do poprzedniego zmienia sytuację na wprost przeciwną: na górnej płytce pojawiają się ładunki dodatnie, na dolnej — ujemne.

Zastosowanie siły ściskającej o kierunku prostopadłym do osi x_1 daje efekt taki, jak rozciąganie w kierunku x_1.

Omówione rysunki przedstawiają schematycznie zmiany zachodzące w pojedynczej komórce krystalicznej złożonej z 3 jonów krzemu i 6 jonów tlenu. Podobne zjawisko, tylko na odpowiednio większą skalę, wystąpi, jeżeli z kryształu kwarcu odpowiednio wytniemy płytkę krystaliczną i poddamy ją również odpowiednio skierowanym siłom rozciągającym lub ściskającym. Sposób wycięcia płytki przedstawia rys. 12.24. Jest to tzw. wykrój X lub wykrój Curie. Poddanie takiej płytki działaniom sił równoległych do osi x lub osi y wywołuje pojawienie się ładunków na powierzchniach prostopadłych do osi x. Działanie sił równoległych do osi z nie daje efektu elektrycznego.



Rys. 12.24

Poddawanie płytki o wykroju Curie działaniu sił periodycznie zmiennych zgniatających i rozciągających, skierowanych np. równolegle do osi x, powoduje periodycznie zmienne pojawianie się ładunków dodatnich i ujemnych na powierzchni płytki. Okazało się dalej, że to zjawisko, zwane piezoelektrycznym, jest odwracalne: jeżeli płytkę kwarcową o wykroju Curie umieścimy w szybko zmiennym polu elektrycznym (między dwiema płytkami kondensatora), wtedy otrzymamy szybko po sobie następujące wydłużenia i skrócenia płytki. Innymi słowy, wystąpią w tych warunkach drgania mechaniczne o dużej częstotliwości. Z wytwarzaniem szybko zmiennych pól elektrycznych nie ma dziś trudności, a zatem poznaliśmy sposób wytwarzania ultradźwięków o bardzo dużej częstotliwości.

Przeanalizujmy bliżej drgania takiej płytki w zmiennym polu elektrycznym. Przede wszystkim może być interesujące, jak wielkie zmiany długości wystąpią w płytkach kwarcowych. Okazuje się, że gdy pole elektryczne jest równoległe do osi x, to płytka o grubości $L_x = 0{,}01$ m wydłuża się o 21,3 Å, gdy różnica potencjałów między okładkami wynosi 1000 V.

Gdy chcemy otrzymywać możliwie duże moce, wykorzystujemy zjawisko rezonansu, tzn. uzgadniamy częstotliwość drgań własnych płytki i częstotliwość zmian pola elektrycznego. Częstotliwość drgań własnych ν_0 dla płytek Curie wyraża się wzorem

$$\nu_0 = \frac{1}{2L_x}\sqrt{\frac{\text{moduł Younga w kierunku osi } x}{\varrho}},$$

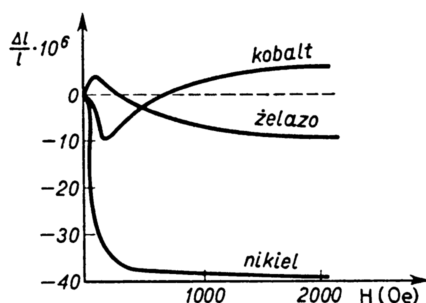
gdzie L_x oznacza grubość płytki w kierunku osi x wyrażoną w m, ϱ — gęstość kwarcu w kg/m^3 (mówimy o module sprężystości na wydłużenie w kierunku osi x ze względu na anizotropię kwarcu).

Wstawienie danych liczbowych do wzoru prowadzi do następującego wyniku:

$$\nu_0 = \frac{284 \cdot 10^5}{L_x} \text{ Hz}.$$

Innymi słowy, częstotliwość drgań własnych płytki jest odwrotnie proporcjonalna do grubości płytki w kierunku osi x. Nie można jednak posuwać się za daleko w kierunku zmniejszania rozmiarów płytki, gdyż płytki bardzo cienkie są za słabe jako źródło drgań ultraakustycznych dla celów przemysłowych. Tak np. płytka o grubości 0,056 mm i o częstotliwości rezonansowej $\nu_0 = 50 \cdot 10^6$ Hz do celów przemysłowych już się nie nadaje. W badaniach laboratoryjnych udało się dojść do częstotliwości 10^9 Hz. Graniczną mocą, przy której można pracować, jest 25 W na każdy centymetr kwadratowy powierzchni płytki.

Dla lepszego przekazania energii drgań otoczeniu zwykle odpowiednio zamocowane płytki kwarcowe umieszcza się w oleju transformatorowym. Ma to jeszcze jedną dobrą stronę: stosunkowo wysoka wartość izolacyjna oleju zmniejsza obawy przebicia podczas stosowania silnych pól elektrycznych między okładkami kondensatora. Przy cienkich płytkach nie można stosować zbyt silnych pól elektrycznych, bo płytka ulega zniszczeniu. Według badaczy radzieckich, gdy grubość płytki waha się od 2–8 mm, to bez awarii można stosować napięcie $1{,}5 \cdot 10^3$ V na mm grubości płytki. Dla częstotliwości rezonansowej około $0{,}5 \cdot 10^6$ Hz przyspieszenie występujące w ruchu drgającej płytki dochodzi do olbrzymiej wartości $22{,}5 \cdot 10^6$ cm/s², prędkość — do 7,3 cm/s, amplituda drgań — do 240 Å. Obecnie stosuje się płytki Curie o różnym kształcie i grubości od 1,3–0,05 mm, co odpowiada częstotliwościom od $220 \cdot 10^3$–$57 \cdot 10^6$ Hz.



Rys. 12.25

Trzeci sposób otrzymywania ultradźwięków polega na wykorzystaniu zjawiska magnetostrykcji. Istotę tego zjawiska stanowi zmiana rozmiarów ferromagnetyka, towarzysząca zmianom natężenia pola magnetycznego.

Zachowanie się różnych ferromagnetyków w zmiennym polu magnetycznym charakteryzuje krzywa (rys. 12.25), przedstawiająca zależność względnych zmian długości $\Delta l/l$ od natężenia pola magnetycznego H. Różne ferromagnetyki zachowują się bardzo różnie. Przeciętne względne zmiany długości np. sztabek ferromagnetycznych są rzędu 10^{-5}–10^{-6}.

Sztabka ferromagnetyczna z nie namagnesowanego np. Ni umieszczona wewnątrz zwojnicy, przez którą płynie prąd zmienny, drga rozszerzając się i kurcząc z często-

tliwością równą podwojonej częstotliwości zmian prądu (kurczenie się i rozszerzanie sztabki towarzyszy każdemu przejściu od wzrostu do zmniejszania się wartości bezwzględnej natężenia pola, a nie zależy od zmiany kierunku).

Do budowy generatorów magnetostrykcyjnych nadają się sztabki z Ni i z niektórych stopów, np. z inwaru (stop Ni 36% i Fe 64%). Mają one odpowiednie własności magnetyczne i wytrzymałościowe. Podobnie jak przy zjawisku piezoelektrycznym, najsilniejsze drgania otrzymujemy stosując częstotliwości rezonansowe. Częstotliwość drgań własnych sztabek jest odwrotnie proporcjonalna do ich długości, a więc znowu do otrzymywania wysokich częstotliwości trzeba stosować krótkie sztabki. Tak na przykład częstotliwości $100 \cdot 10^3$ Hz odpowiada długość sztabki ok. 2,5 cm. Właściwie w praktyce krótszych sztabek się nie stosuje ze względu na trudność ich zamocowania i odpowiedniego zawieszenia wewnątrz cewki wytwarzającej zmienne pole magnetyczne. Ponadto natężenie otrzymywanych drgań maleje w miarę wzrostu częstotliwości, dlatego też najczęściej nadajniki magnetostrykcyjne pracują na częstotliwości 70 tysięcy herców.

Trzeba pamiętać, że efekt magnetostrykcyjny w dużym stopniu zależy od temperatury. Ze wzrostem temperatury efekt ten maleje spadając do zera w tzw. temperaturze Curie, przy której ciała tracą właściwości ferromagnetyczne.

Przechodzimy obecnie do omówienia zastosowań ultradźwięków. Są one bardzo różnorodne. Fale ultradźwiękowe w cieczach mogą wywoływać powstawanie emulsji, np. rtęci w wodzie, oliwy w wodzie, AgCl lub AgBr w żelatynie. Praktycznie jest to wykorzystane przy produkcji drobnoziarnistych emulsji fotograficznych oraz w przemyśle farmaceutycznym przy produkcji kremów i lekarstw.

Wysoko spolimeryzowane cząstki pod wpływem fal ultradźwiękowych mogą ulegać rozpadowi, np. krochmal rozpada się na dekstrynę.

Znane jest także działanie koagulujące fal ultradźwiękowych: drobne cząstki zawiesin (hydrozole i aerozole) mogą się łączyć w większe zespoły. Z tym wiąże się zagadnienie usuwania mgły: połączenie drobnych kropelek mgły w większe krople sprzyja opadaniu ich na ziemię. Analogicznie łączenie się drobnych cząstek sadzy w większe agregaty pod wpływem fal ultradźwiękowych w długim przewodzie kominowym sprzyja ich opadaniu wewnątrz komina i przyczynia się do oczyszczania powietrza wielkich miast fabrycznych.

Pod działaniem ultradźwięków giną bakterie i niższe zwierzęta. Czerwone ciałka krwi ulegają rozkładowi. Ultradźwięki są pomocą dla lekarza. Oddają cenne usługi przy leczeniu schorzeń reumatycznych, a także w pewnych przypadkach zastępują nóż chirurga. W tym celu uzyskuje się zbieżne wiązki fal ultradźwiękowych przez przepuszczenie ultradźwięków przez odpowiednie soczewki skupiające z mas plastycznych. Wiązki te kieruje się tak, aby ich punkt zbieżności (czyli miejsce maksymalnego natężenia) przypadał np. w ognisku jakiegoś wewnętrznego ropnia. Działanie bakteriobójcze ultradźwięków przyczynia się do zlikwidowania ropnia bez zabiegu chirurgicznego.

Fale ultradźwiękowe znalazły szerokie zastosowanie w metalurgii. Wymienimy tylko parę ważniejszych przykładów. Poddanie roztopionych metali działaniu ultra-

dźwięków w temperaturze krzepnięcia powoduje zmianę ich struktury mikrokrystalicznej, decydującej m.in. o wytrzymałości, kruchości i innych cechach metali. Równocześnie obniża się temperatura krzepnięcia i wzrasta szybkość zastygania. Ponadto ultradźwięki są wykorzystywane do prześwietlania bloków czy odlewów metalicznych celem wykrywania wad na głębokościach kilku metrów, robiąc w tej dziedzinie poważną konkurencję promieniom Röntgena. Jest to tzw. defektoskopia ultradźwiękowa. Gdy wchodzą w grę grube bloki metalowe, promienie Röntgena nie spełniają swego zadania.

W przemyśle chemicznym ultradźwięki mogą służyć m. in. do odgazowywania cieczy. Prześwietlenie cieczy ultradźwiękami powoduje szybkie usunięcie rozpuszczonych w niej gazów. To samo zresztą może być stosowane w przemyśle metalurgicznym, gdyż jak wiadomo, pęcherzyki gazów stanowią najpospolitszą wadę odlewów.

Dalszym ważnym zastosowaniem ultradźwięków jest hydrolokacja ultradźwiękowa. Hydrolokacja polega na wysyłaniu przez generator znajdujący się na statku krótkotrwałego impulsu ultradźwiękowego w stronę dna morskiego. Impuls taki odbija się od dna i wraca jako echo ultraakustyczne, wywołując w aparacie odbiorczym krótkotrwałe drganie płytki kwarcowej. Drganiom płytki towarzyszy pojawienie się zmiennych ładunków elektrycznych na jej powierzchni, co wywołuje odpowiednie drganie w obwodzie elektrycznym i może być łatwo zarejestrowane. Mierząc czas wysłania impulsu i czas jego powrotu można określić głębokość morza, znając prędkość rozchodzenia się ultradźwięków w wodzie.

Metoda impulsów ultraakustycznych stosowana w hydrolokacji, służąca początkowo do ustalania głębokości dna morskiego, wykrywania łodzi podwodnych, gór lodowych i skał, została następnie rozszerzona na impulsy fal elektromagnetycznych, wykorzystywane w urządzeniach radarowych przy tzw. radiolokacji.

Prędkości fal dźwiękowych w cieczach są rzędu 1000 m/s. Jedna z metod pomiaru prędkości fali ultradźwiękowej polega na wytworzeniu w cieczy fali ultraakustycznej zgęszczeń i rozrzedzeń. Dla wiązki światła padającej z boku te zgęszczenia i rozrzedzenia odgrywają podobną rolę jak rysy w siatce dyfrakcyjnej (por. rozdz. 25 o interferencji światła). Gdy zastosujemy odpowiedni układ optyczny, te zgęszczenia i rozrzedzenia spowodują wystąpienie szeregu prążków dyfrakcyjnych. Mierząc odległość prążków łatwo można ustalić długość fali ultradźwiękowej w cieczy, a następnie z zależności $\lambda = v/\nu$ obliczyć prędkość fali ultradźwiękowej.

Liczne badania prędkości fal ultradźwiękowych w cieczach nie doprowadziły jeszcze do znalezienia zależności między prędkością a składem chemicznym i budową cząsteczek cieczy. Znalezienie takiej zależności dostarczyłoby chemikom nowego, bardzo cennego narzędzia badań.

Pytania i zadania

1. Podać ogólną charakterystykę fali głosowej uwzględniając: a) rodzaj fali, b) prędkość rozchodzenia się fali w dowolnym ośrodku sprężystym, c) prędkość fali głosowej w ośrodku gazowym.

2. Wymienić i zdefiniować znamiona dźwięków.

3. Omówić wrażliwość ucha ludzkiego na tony o różnych częstotliwościach na podstawie krzywej czułości ucha ludzkiego.

4. Zdefiniować: a) dolny próg słyszalności i b) próg bólu dla tonu o określonej częstotliwości.

5. Czy próg słyszalności jest synonimem dolnej granicy słyszalności?

6. Jaka jest definicja poziomu ciśnienia akustycznego i poziomu głośności? W jakich jednostkach wyrażony jest poziom ciśnienia akustycznego, poziom głośności i natężenie dźwięku?

7. Podać krótką charakterystykę takich zjawisk jak: a) odbicie fali głosowej i powstawanie echa, b) załamanie fali głosowej, c) ugięcie fali głosowej.

8. Wyjaśnić związek efektu dudnień z interferencją fal.

9. W jakich warunkach interferencja fal głosowych prowadzi do wytwarzania się fal stojących?

10. Omówić kilka przykładów tworzenia się fal stojących akustycznych.

11. W jakich warunkach powstaje zjawisko Dopplera? Jakie rozumowanie doprowadza do wzoru na częstotliwość drgań odbieranych? Czy częstotliwość drgań odbieranych zmienia się w sposób ciągły przy ruchu jednostajnym źródła lub obserwatora (tzn. czy obserwator odbiera ton coraz wyższy lub coraz niższy)?

12. Dlaczego gdy porusza się tylko źródło lub tylko obserwator to, nawet przy jednakowej wartości liczbowej ich prędkości względem nieruchomego ośrodka, obserwuje się różne zmiany częstotliwości tonu odbieranego względem nadawanego?

13. Podać ogólną charakterystykę ultradźwięków.

14. Omówić metodę mechaniczną, piezoelektryczną i magnetostrykcyjną otrzymywania ultradźwięków.

15. Omówić podstawowe zastosowania ultradźwięków.

16. Wykazać, że prędkość głosu w powietrzu w temperaturze TK można wyrazić przybliżonym wzorem

$$v = 20\sqrt{T}\ \mathrm{m\cdot s^{-1}}.$$

Współczynnik $\varkappa$ dla powietrza przyjąć 1,4, gęstość powietrza w warunkach normalnych 1,3 $\mathrm{kg\cdot m^{-3}}$.

17. Natężenie pewnego dźwięku wzrosło pięciokrotnie. O ile wzrósł przy tym poziom ciśnienia akustycznego?

Odp. ok. 7 decybeli.

18. Poziom ciśnienia akustycznego pewnego dźwięku wzrósł o jeden decybel. W jakim stosunku powiększyło się natężenie tego dźwięku?

Odp. 1,26.

19. Obliczyć siłę naciągającą strunę o długości 0,5 m, jeżeli wiadomo, że częstotliwość jej drgań podstawowych wynosi 20 Hz, a masa jednostki długości 1,5 $\mathrm{g\cdot cm^{-1}}$.

Odp. 6 kG.

20. Dwie piszczałki jednostronnie zamknięte o jednakowej długości $L = 60$ cm dają jedno dudnienie na sekundę w związku z różnicą temperatur powietrza w obu piszczałkach. Wiedząc, że temperatura powietrza w piszczałce wydającej ton niższy wynosi 16°C wyznaczyć temperaturę powietrza w drugiej piszczałce.

Odp. 20°C.

21. Częstotliwość tonu wydawanego przez parowóz zbliżający się do obserwatora z prędkością 54 km/h wynosi 650 Hz. Jaką częstotliwość będzie miał ton odbierany przez obserwatora, jeżeli temperatura powietrza wynosi 16°C?

Odp. 680 Hz.

22. Nieruchomy obserwator stwierdza, że stosunek odbieranych przez niego częstotliwości dźwięku syreny samochodu przy zbliżaniu się i następnie oddalaniu się samochodu ruchem jednostajnym wynosi $\frac{9}{8}$. Zakładając, że powietrze jest nieruchome i że prędkość głosu w powietrzu wynosi 340 m/s, obliczyć prędkość samochodu.

Odp. 72 km/h.

23. Drgający kamerton znajdujący się między obserwatorem a ścianą przesuwa się prostopadle w stronę ściany z prędkością 25 cm/s. Nieruchomy obserwator odbiera trzy dudnienia w czasie 1 sekundy. Wyznaczyć częstotliwość drgań kamertonu zakładając, że ośrodek jest nieruchomy i że prędkość głosu w powietrzu równa się 340 m/s.

Odp. 2040 Hz.

CZĘŚĆ II Ciepło*

ROZDZIAŁ 13

Termometria

13.1. Podstawy termometrii

Jednym z podstawowych pojęć fizycznych jest pojęcie temperatury. Molekularno-kinetyczne określenie temperatury poznamy w § 16.6. Do tej chwili będziemy traktowali temperaturę jako wielkość charakteryzującą stopień ogrzania danego ciała. Musimy przede wszystkim ustalić:

a) w jakich warunkach będziemy uważali temperatury kilku ciał za jednakowe,

b) kiedy temperaturę dowolnego ciała A uznamy za wyższą od temperatury ciała B,

c) kiedy temperaturę dowolnego ciała A uznamy za niższą od temperatury ciała B.

Prawem doświadczalnym, na którym opiera się możliwość pomiaru temperatury, jest prawo wyrównywania się temperatur ciał nie odizolowanych od siebie cieplnie, jeżeli żadne z nich nie jest źródłem energii cieplnej. Mówimy, że takie ciała pozostające w zetknięciu ze sobą przez czas dłuższy są w *stanie równowagi termicznej*: energia cieplna przechodzi od jednego ciała do drugiego w jednakowej ilości; *temperatury ciał są jednakowe.*

Jeżeli ciała A i B pozostają w zetknięciu i energia cieplna przechodzi w nadmiarze od ciała A do ciała B, mówimy, że ciało A ma temperaturę wyższą niż B. Odwrotnie, gdy energia cieplna przechodzi w nadmiarze od ciała B do A, temperatura ciała A jest niższa od temperatury ciała B. Ustalenie pojęcia równości i nierówności temperatur jeszcze nie określa sposobu porównywania temperatur różnych ciał. Do tego celu służą termometry.

Pomiar temperatury dowolnego ciała za pomocą termometru opiera się na prawie wyrównywania się temperatur. Termometr wskazuje zawsze temperaturę własną, ale

* Zagadnienia związane z ciepłem wymagają pewnych uwag ogólnych. Ciepło jest to pojęcie znane już Czytelnikowi z kursu szkoły średniej, gdzie było definiowane jako wielkość równoważna pracy, jako rodzaj energii (energia cieplna). Według nowszych poglądów ciepło — podobnie jak praca — powinno być raczej traktowane jako postać (forma) przekazywania energii. Obecnie jednak w nowym ujęciu wyłania się jeszcze wiele nieuzgodnionych spraw terminologicznych. W związku z tym — dla uniknięcia niepotrzebnego zamieszania — do chwili ustalenia obowiązującej, nowej terminologii będziemy jeszcze traktowali ciepło jako rodzaj energii.

dzięki temu, że przez pewien czas pozostaje w zetknięciu z ciałem badanym, mamy prawo wnioskować, że osiągnięty jest stan równowagi termicznej.

W termometrach wykorzystywane są: a) różne substancje termometryczne, b) różne właściwości termometryczne ciał. W zależności od zastosowanej substancji termometrycznej można wyróżnić termometry rtęciowe, alkoholowe, pentanowe, platynowe itp. Jako właściwość termometryczna może być wybrana dowolna wielkość charakteryzująca stan ciała termometrycznego, zmieniająca się w sposób ściśle określony, jednoznaczny przy zmianach temperatury. Tak np. objętość rtęci stale wzrastająca ze wzrostem temperatury wykorzystywana jest jako właściwość termometryczna w termometrze rtęciowym. Prężność gazu zmieniająca się przy stałej objętości proporcjonalnie do zmian temperatury wykorzystywana jest jako właściwość termometryczna w termometrze gazowym. Właściwością termometryczną może też być opór elektryczny, ulegający zmianom podczas zmian temperatury.

Zmiana objętości wody podczas ogrzewania, np. w granicach od 0 do 10°C, nie byłaby odpowiednią właściwością termometryczną, gdyż dwom różnym temperaturom (jednej niższej od 4°C, drugiej od 4°C wyższej) odpowiadałaby ta sama objętość cieczy termometrycznej.

W poprzednim zdaniu posłużyliśmy się już pewnymi — znanymi zresztą dobrze z życia codziennego — jednostkami służącymi do wyrażania temperatur, a mianowicie stopniami Celsjusza. Sprawa skal termometrycznych wymaga jednak kilku ogólnych uwag dodatkowych.

W układzie SI obowiązującą skalą temperatur jest tzw. skala termodynamiczna, wprowadzona w oparciu o prawa termodynamiki. Jej podstawy teoretyczne omówimy w rozdz. 18. Obecnie poprzestaniemy na stwierdzeniu, że w tej skali (którą krótko będziemy czasem nazywali skalą Kelvina) jako punkt zerowy przyjmuje się *temperaturę zera bezwzględnego*, a jako drugi punkt odniesienia — *temperaturę* tzw. *punktu potrójnego wody** określoną liczbą 273,16. Te umowy, dotyczące temperatur odniesienia, definiują jednostkę temperatury w skali termodynamicznej, zwaną kelwinem (K).

Temperaturę wyrażoną w skali Kelvina oznaczamy symbolem T. Skalą temperatur przejściowo dopuszczoną do stosowania jest skala Celsjusza. Jednostką w tej skali jest stopień Celsjusza (°C), a symbolem wyrażającym ogólnie temperaturę w tej skali jest t. Związek skali Celsjusza ze skalą termodynamiczną ma następującą postać:

$$T = t+273,15$$

(temperatura punktu potrójnego wody w skali Celsjusza wynosi 0,01°C). Innymi słowy, początki obu skal są przesunięte względem siebie o 273,15 stopni, ale przyrosty temperatur w obu skalach wyrażają się tą samą liczbą stopni. A zatem w odniesieniu do przyrostów temperatur można stosować wspólną jednostkę kelwin (K).

---

* Por. § 3 „Wiadomości wstępnych".

Umowy międzynarodowe ustaliły pewne podstawowe temperatury odniesienia. Do nich należą m.in.: temperatura punktu potrójnego wody (0,01°C) oraz mierzone pod normalnym ciśnieniem temperatury wrzenia tlenu (—182,97°C), wody (100°C), siarki (444,60°C) i temperatury topnienia lodu (0,00°C), srebra (960,80°C), złota (1063,00°C).

13.2. Termometry cieczowe

Z tej klasy termometrów najbardziej rozpowszechnione są termometry rtęciowe. Podczas cechowania termometru rtęciowego odstęp między poziomami rtęci, odpowiadającymi temperaturom 0°C i 100°C, dzielimy na sto równych części i następnie zachowując wielkość podziałki stopniowej rozszerzamy skalę w stronę temperatur wyższych i niższych. Rurka włoskowata termometru musi mieć na całej długości niezmienny przekrój poprzeczny.

Różne termometry rtęciowe, mimo że jednakowo wskazują temperatury 0 i 100, mogą dawać różne wskazania w innych temperaturach. Wynika to stąd, że podczas cechowania brane są pod uwagę tylko dwa punkty podstawowe i zakładana jest równomierna rozszerzalność rtęci i szkła w całym szerokim zakresie stosowalności termometru rtęciowego. Tymczasem w tych szerokich granicach temperatur różne gatunki szkła zachowują się różnie.

Większe jeszcze różnice wskazań występują, gdy do pomiaru temperatury stosujemy inne termometry, nawet wtedy, gdy ich działanie oparte jest na tej samej właściwości termometrycznej, tzn. na rozszerzalności cieplnej cieczy. Termometry alkoholowe lub pentanowe o równomiernej podziałce dają wskazania zgodne z termometrem rtęciowym, gdy wchodzą w grę temperatury wykorzystane podczas cechowania. Dla innych temperatur występują większe lub mniejsze odchylenia w związku z nierównomierną rozszerzalnością cieczy termometrycznych w granicach temperatur wchodzących w rachubę. Można uniknąć tych rozbieżności wprowadzając w termometrach niertęciowych nierównomierne podziałki i cechując je przez dokładne porównywanie z termometrem rtęciowym w licznych temperaturach.

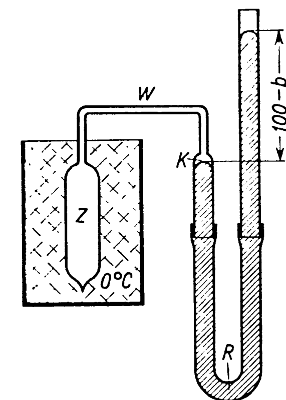
Zakres stosowalności różnych rodzajów termometrów cieczowych jest różny. Termometry rtęciowe mogą być używane w zakresie temperatur około 300°C, a mianowicie od około —30°C (temperatura krzepnięcia rtęci pod normalnym ciśnieniem wynosi —38,87°C) do około +300°C, gdyż temperatura wrzenia rtęci pod normalnym ciśnieniem wynosi 357°C. Górną granicę stosowalności termometru rtęciowego można przesunąć aż do 800°C przez wypełnienie wolnej przestrzeni nad rtęcią w rurce włoskowatej azotem pod zwiększonym ciśnieniem (do ok. $80 \cdot 10^5$ N/m^2, czyli do 80 atm). Wiąże się to z wpływem zwiększonego ciśnienia na temperaturę wrzenia cieczy. W takich termometrach przeznaczonych do mierzenia wysokich temperatur stosujemy specjalne gatunki szkła lub też szkło zastępujemy kwarcem. Do mierzenia temperatur niskich nadają się termometry alkoholowe do —100°C i pentanowe do —190°C.

13.3. Termometry gazowe

Jako wzorcowy termometr służący do cechowania i sprawdzania termometrów cieczowych używany jest termometr gazowy. Substancją termometryczną w termometrze gazowym idealnym powinien być gaz doskonały. Ponieważ jednak taki gaz nie istnieje, więc termometry gazowe wypełnia się wodorem, helem lub azotem, jako gazami najbardziej zbliżonymi właściwościami do gazu doskonałego.

Na rysunku 13.1 pokazana jest zasada budowy termometru wodorowego. Zbiornik Z przechodzi w rurkę włoskowatą W, do której dołączony jest manometr rtęciowy. Oba ramiona manometru połączone są ze sobą rurką gumową R, umożliwiającą dowolne przesuwanie ramienia otwartego. Zbiornik Z oraz cała rurka włoskowata wypełnione są wodorem.

Pomiar temperatury za pomocą termometru gazowego opiera się na badaniu zmian prężności gazu ogrzewanego w stałej objętości. Zgodnie z jednym z praw rządzących gazami doskonałymi (por. § 14.4) *zmiany prężności gazu podczas ogrzewania w stałej objętości są wprost proporcjonalne do przyrostów temperatury.* Podczas ogrzewania wodoru zawartego w zbiorniku Z zmienia się objętość gazu i jego prężność. Rozszerzający się gaz wypycha rtęć w stronę otwartego ramienia manometru. Zmieniając położenie tego ramienia sprowadzamy wodór do objętości pierwotnej określonej położeniem zakończenia kolca K. Gdy powierzchnia rtęci styka się z kolcem, wodór zajmuje objętość V_0. Umowa międzynarodowa normuje ilość wodoru wprowadzonego do zbiornika. Masa wodoru w termometrze gazowym powinna być taka, aby w temperaturze 0°C zajmując objętość V_0 wywierał on ciśnienie 1 metra słupa rtęci. Innymi słowy, gdy zbiornik Z umieszczony jest w kąpieli z topniejącym lodem, a ramię otwarte tak ustawione, że poziom rtęci w ramieniu od strony zbiornika styka się z kolcem K, to różnica poziomów rtęci w obu ramionach manometru musi być równa $(100-b)$ cm słupa rtęci. Przez b oznaczamy ciśnienie atmosferyczne wskazywane przez barometr w chwili pomiaru temperatury.



Rys. 13.1

Pomiar temperatury wymaga wycechowania termometru. W tym celu zanurzamy zbiornik Z do kąpieli z topniejącym lodem i przesuwamy prawe ramię tak, aby rtęć w lewym ramieniu dotykała kolca K (wtedy objętość wodoru jest równa objętości początkowej V_0). Następnie odczytujemy różnicę poziomów w obu ramionach i ustalamy wartość ciśnienia wodoru p_0 w temperaturze 0°C. W analogiczny sposób ustalamy wartość ciśnienia p_{100} po wstawieniu zbiornika Z do pary nad wodą wrzącą pod normalnym ciśnieniem.

Zmianie temperatury o 100°C odpowiada przyrost ciśnienia $p_{100}-p_0$, a więc zmianie temperatury o 1°C odpowiada przyrost ciśnienia

$$\frac{p_{100}-p_0}{100}.$$

Po zanurzeniu zbiornika do kąpieli o dowolnej temperaturze t ustalamy wartość p_t przy stałej objętości V_0. Na podstawie proporcjonalności przyrostów prężności do przyrostów temperatury przyrostowi ciśnienia $p_t - p_0$ odpowiada przyrost temperatury od 0 do t, czyli

$$t = \frac{p_t - p_0}{\frac{p_{100} - p_0}{100}} = \frac{p_t - p_0}{p_{100} - p_0} 100.$$

Posługując się termometrem wodorowym uwzględniamy poprawki związane z rozszerzalnością zbiornika Z i z objętością rurki włoskowatej W, która znajduje się w innej temperaturze niż cały zbiornik. Objętość tej rurki V_1 jest zresztą bardzo mała w porównaniu z objętością całego zbiornika, gdyż zazwyczaj

$$\frac{V_1}{V} < 0{,}001.$$

Zakres stosowalności termometru wodorowego jest szeroki. Sięga od -240 do $+300°C$. W temperaturach wyższych wodór przenika przez ścianki naczynia. Do mierzenia temperatur wyższych, do 1600°C, służy termometr azotowy. Zbiornik szklany zastąpiony jest wtedy zbiornikiem kwarcowym lub porcelanowym. Do mierzenia temperatur niższych od $-200°C$ służą termometry gazowe wypełnione helem.

Odchylenia wskazań termometru azotowego od wskazań termometru gazowego idealnego są większe niż w termometrze wodorowym lub helowym. Odchylenia te dla termometrów wodorowego, helowego i azotowego są zestawione w tabelkach, które umożliwiają przeliczanie wskazań termometrów gazowych na wskazania termometru gazowego idealnego.

13.4. Termometry elektryczne

Prócz wymienionych sposobów pomiaru temperatur stosuje się jeszcze metody optyczne i elektryczne. Metody optyczne omówimy szerzej w kursie optyki. Metody elektryczne wykorzystują jako właściwość termometryczną zależność temperaturową oporu elektrycznego albo siły elektromotorycznej termoogniw (termopar).

Termometry oporowe platynowe stosowane są w granicach temperatur -200 i $+550°C$. Opór platyny w granicach temperatur od 0–550°C zmienia się według równania

$$R_t = R_0(1 + \alpha t + \beta t^2),$$

gdzie R_t oznacza opór platyny w temperaturze badanej t, R_0 — opór platyny w temperaturze 0°C, α i β — współczynniki stałe. Przed przystąpieniem do zasadniczego pomiaru wyznacza się doświadczalnie wielkość R_0 oraz oblicza się α i β po zmierzeniu oporu R w dwóch znanych temperaturach, np. w 100°C i w temperaturze wrzenia siarki 444,55°C.

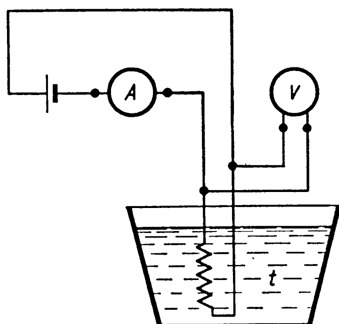
Rozszerzając zakres stosowalności tego termometru do około $-200°C$ musimy już opierać się na bardziej złożonej zależności oporu od temperatury, a mianowicie

$$R_t = R_0(1 + ct + dt^2 + et^3).$$

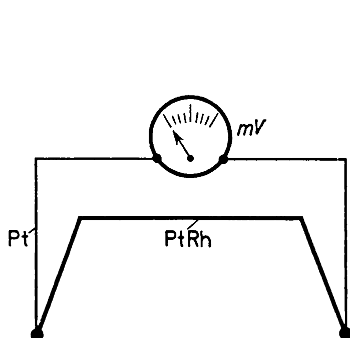
Stałe c, d, e, występujące w tym wzorze, wyznacza się podobnie jak α i β przez zmierzenie oporu w trzech temperaturach ściśle określonych, np. w temperaturze krzepnięcia rtęci —38,87°C, wrzenia dwutlenku węgla —78,51°C i wrzenia tlenu —183,0°C pod ciśnieniem normalnym.

Różnice wskazań termometru oporowego platynowego i termometru gazowego w granicach 0–100°C nie przekraczają 0,01°C.

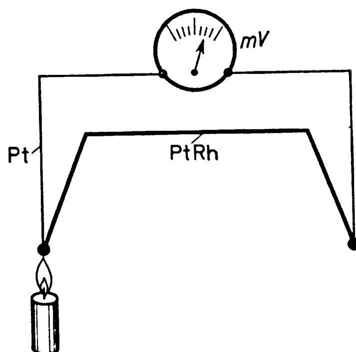
Rysunek 13.2 przedstawia schemat połączeń przy stosowaniu termometru oporowego.



Rys. 13.2



Rys. 13.3



Rys. 13.4

W granicach 600–1200°C (a w technice i nieco wyżej) używane są termoogniwa z platyny i platynorodu (stop platyny z 10% rodu). Często też do mierzenia niższych temperatur stosuje się termoogniwa z żelaza i konstantanu (do +600°C), miedzi i konstantanu (do —200°C), niklu i chromoniklu (do +1100°C) itp.

Utwórzmy zamknięty obwód elektryczny z drutu platynowego i platynorodowego (rys. 13.3) i oba miejsca spojenia pozostawmy w tej samej temperaturze. Prąd elektryczny w obwodzie nie popłynie. Włączony do obwodu miliwoltomierz nie wykaże żadnej siły elektromotorycznej. Wystarczy jednak zmienić temperaturę jednego ze spojeń

(rys. 13.4), aby wskazówka miliwoltomierza wykazała istnienie siły elektromotorycznej, zwanej termoelektryczną. Zjawisko termoelektryczne będzie omówione szerzej w kursie elektryczności. Wartość siły termoelektrycznej w danym termoogniwie zależy od różnicy temperatur spojeń. Utrzymując jedno spojenie w temperaturze stałej można tak wycechować wskazania miliwoltomierza, aby bezpośrednio określały one temperaturę drugiego spojenia.

We wspomnianym wyżej zakresie temperatur 600–1200°C wskazania tego termometru elektrycznego nie różnią się więcej niż o 1°C od wyników otrzymywanych za pomocą termometru gazowego. Powyżej 1200°C rozbieżności sięgają już 5°C.

Przy stosowaniu termoogniw do pomiaru temperatury należy pilnie przestrzegać unikania takich warunków, które mogą powodować uszkodzenie przyrządu. Uszkodzenia mogą być typu:

a) chemicznego — oddziaływanie badanego środowiska na materiał termoogniwa,

b) mechanicznego — np. odkształcenia sprężyste,

c) termicznego — przegrzanie termoogniwa.

Przy długotrwałym użyciu właściwości termoogniwa mogą ulegać zmianie wskutek procesu starzenia materiału termoogniwa.

Niedokładność pomiaru temperatury może wynikać z uszkodzeń miliwoltomierza wywołanych działaniem żrących par lub gazów (konieczne jest stosowanie szczelnej obudowy), działaniem wysokiej temperatury (unikanie przegrzań) lub też drganiem maszyn znajdujących się w bliskim otoczeniu (stosowanie amortyzujących podkładek z gumy).

Jeżeli druty w pobliżu spojeń mają średnice bardzo niewielkie, to całe spojenie wraz z najbliższym otoczeniem przypomina cienką igłę. Takie urządzenie można stosować do dokładnego pomiaru temperatury bardzo małych obiektów. W mikrotermoogniwie spajane druciki mają średnicę około 0,02 mm, są więc tak cienkie jak najcieńszy włos ludzki. Sporządzenie i spojenie takich drucików jest już samo w sobie pewnym problemem technicznym.

Pomiar temperatury za pomocą termoogniw, jak również za pomocą termometrów oporowych może być bardzo dokładny, jeżeli zastosujemy wysokiej klasy mierniki elektryczne. Trzeba pamiętać, że zmiany oporu metali w zależności od temperatury są nieznaczne. Tak np. wzrost temperatury o 300°C powoduje dwukrotny wzrost oporu platyny, a trzykrotny wzrost oporu żelaza.

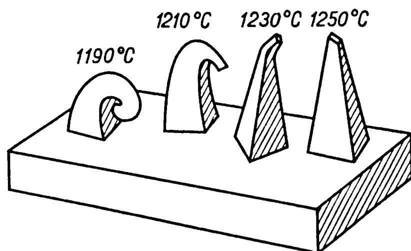
Ostatnio opracowano nową metodę pomiaru temperatury, opartą na zależności oporu półprzewodników od temperatury. Półprzewodniki są to ciała, których opory właściwe zawarte są w granicach 1–10^8 $\Omega \cdot$ m. Do tej kategorii należą przede wszystkim tlenki i siarczki metali. Opór tych ciał ze wzrostem temperatury na ogół maleje w przeciwieństwie do oporu metali. Różnica temperatur około 300°C wywołać może np. 1000-krotną zmianę oporu półprzewodnika. Odpowiednio spreparowane półprzewodniki, np. w postaci maleńkich kulek z doprowadzonymi do nich cieniutkimi elektrodami, stanowią po włączeniu do obwodu elektrycznego tzw. *termometry termistorowe*. Zetknięcie takiej kulki z półprzewodnika z ciałem badanym, np. z ciałem ludzkim, pozwala w ciągu 1 s (a nie 10 minut — jak przy stosowaniu zwykłego termometru le-

karskiego) wyznaczyć temperaturę na odpowiednio wycechowanym przyrządzie elektrycznym włączonym do obwodu.

Termistory mogą służyć do pomiaru temperatur przepływających cieczy lub gazów, można się też nimi posługiwać do mierzenia temperatur np. różnych warstw zboża w magazynach, czy też temperatur części kotłów i maszyn trudno dostępnych przy użyciu większych przyrządów. Meteorologiczne termometry termistorowe włączone do odpowiednich obwodów na balonach-sondach automatycznie przekazują na Ziemię drogą radiową wiadomości o temperaturze powietrza na różnych wysokościach. Zastosowania termistorów stają się coraz szersze.

13.5. Uwagi ogólne

Na zakończenie rozdziału poświęconego temperaturze i jej pomiarom dorzucimy parę słów o kilku stosowanych w technice sposobach przybliżonego określania temperatur. Tak np. w przemyśle ceramicznym stosowane są tzw. stożki Segera. Są to zestawy ostrosłupów z mas ceramicznych tak dobrane, że każdy następny zaczyna mięknąć w temperaturze o 10–20 stopni wyższej. Ogólny zakres temperatur objętych takimi stożkami sięga od 600 do 2000 stopni. Zestaw 5–6 takich stożków odpowiednio dobranych do celu, jakiemu ma służyć, osadza się na płycie szamotowej i umieszcza w piecu. Z wyglądu stożków (rys. 13.5) po pewnym czasie można wnioskować o wysokości temperatury. Tak np. w badanym przypadku maksymalna temperatura pieca zawarta jest między 1230 i 1250°C.



Rys. 13.5

Do szacowania temperatury służą też różne termofarby i termokredki zmieniające barwę po osiągnięciu pewnego stopnia ogrzania. Są to najczęściej złożone sole kobaltu, chromu, molibdenu i innych metali. Farby termometryczne mogą być odwracalne i nieodwracalne. Po oziębieniu pierwsze wracają do barwy pierwotnej, drugie zachowują barwę zmienioną.

Termofarby służą najczęściej jako wskaźniki osiągnięcia lub przekroczenia pewnej temperatury o jakimś specjalnym znaczeniu. Tak np. można odpowiednią termofarbą pokryć głowicę silnika spalinowego. Zmiana barwy wskazuje na przegrzanie silnika i jest sygnałem do zastosowania odpowiednich środków ostrożności.

Pytania i zadania

1. Co rozumiemy przez stan równowagi termicznej?
2. Kiedy mówimy o: a) równości, b) nierówności temperatur ciał A i B?
3. Jakie podstawy klasyfikacji można przyjąć dla różnych rodzajów termometrów? Podać przykłady.
4. Jak się przedstawia podział termometrów w zależności od zakresu mierzonych temperatur?
5. Omówić szczegółowo budowę i zasadę działania termometru gazowego.
6. Omówić podział termometrów w zależności od: a) właściwości termometrycznych, b) substancji termometrycznych. Uwzględnić odpowiednie przykłady.
7. Omówić budowę, zasadę działania oraz zastosowania termistorów.
8. Jakie są metody przybliżonej oceny temperatur w przemyśle?
9. Wodór zawarty w naczyniu o stałej objętości wywiera w temperaturze 0°C ciśnienie 1 metra słupa rtęci. Po zanurzeniu tego naczynia w gorącej kąpieli ciśnienie wodoru wzrosło o 300 mm Hg. Wiedząc, że współczynnik prężności wodoru wynosi $\frac{1}{273}$ K^{-1} obliczyć temperaturę kąpieli.

Odp. 81,9°C.

10. Wyznaczyć, dla jakich temperatur mierzonych w skali Celsjusza wskazania termometru wodnego byłyby jednakowe w przedziale od 0 do 10°C, jeżeli wiadomo, że w tym przedziale temperatur objętość wody wyraża się następującym wzorem:

$$V = V_0(1+at+bt^2),$$

gdzie V i V_0 oznaczają odpowiednio objętość wody w temperaturze t i w temperaturze 0°C, $a = 6{,}1 \cdot 10^{-5}$ K^{-1}, $b = -7{,}7 \cdot 10^{-6}$ K^{-2}.

Odp. Dla każdej pary temperatur o wartościach t i $(7{,}9-t)$.

ROZDZIAŁ 14

Rozszerzalność cieplna ciał

14.1. Rozszerzalność liniowa

Zajmiemy się obecnie rozszerzalnością liniową ciał stałych. Oznaczmy długość początkową ciała w temperaturze t_0 przez l_0, długość końcową ciała w temperaturze t przez l. Jeśli różnica temperatur $\Delta t = t - t_0$ nie jest duża (rzędu kilkudziesięciu stopni), to można w pierwszym przybliżeniu przyjmować, że przyrost długości $\Delta l = l - l_0$ jest proporcjonalny do przyrostu temperatury i do długości początkowej:

$$l - l_0 = \bar{\alpha} l_0 \Delta t. \tag{14.1}$$

Przekształcając wzór (14.1) znajdujemy wyrażenie określające długość końcową po ogrzaniu o Δt:

$$l = l_0(1 + \bar{\alpha}\Delta t). \tag{14.2}$$

Współczynnik proporcjonalności $\bar{\alpha}$ ze wzoru (14.1) nosi nazwę *średniego współczynnika rozszerzalności liniowej w granicach temperatur* od t_0 do t:

$$\bar{\alpha} = \frac{l - l_0}{l_0 \Delta t}.$$

Ułamek ten oznacza, o *jaką część długości początkowej wzrasta średnio długość danego ciała po ogrzaniu* o *1 kelwin w granicach temperatur* od t_0 do t. Średni współczynnik rozszerzalności liniowej wyraża się w K^{-1}.

Przechodząc do granicy, gdy Δt zmierza do zera, znajdziemy współczynnik rozszerzalności liniowej odpowiadający danej temperaturze t_0

$$\alpha = \lim_{\Delta t \to 0} \bar{\alpha} = \lim_{\Delta t \to 0} \frac{l - l_0}{l_0 \Delta t},$$

skąd

$$\alpha = \frac{1}{l_0}\frac{dl}{dt}. \tag{14.3}$$

W tabeli 14.1 podane są wartości średnich współczynników rozszerzalności liniowej w granicach temperatur 273–373K. Wymieniony w tabeli kwarc jest przykładem ciała anizotropowego, tzn. ciała mającego w różnych kierunkach różne właściwości.

Tabela 14.1

Średnie współczynniki rozszerzalności liniowej w granicach temperatur 273–373K

Rodzaj ciała	$\bar{\alpha} \cdot 10^6 K^{-1}$	Rodzaj ciała	$\bar{\alpha} \cdot 10^6 K^{-1}$
ołów	29	porcelana	3,5
glin	24	kwarc topiony	0,4
srebro	19	inwar (64% Fe, 36% Ni)	0,4
miedź	17	kwarc krystaliczny	
nikiel	14	(prostopadle do osi)	14
złoto	14	kwarc krystaliczny	
żelazo, stal	10–12	(równolegle do osi)	8
platyna	9		
szkło sodowe	9		
szkło „pyrex"	3		
wolfram	4		

Niektóre ciała podczas ogrzewania kurczą się, czyli mają ujemne współczynniki rozszerzalności liniowej. Do takich należą m. in. rozciągnięty kauczuk, jodek srebra w granicach temperatur od -60 do $+142°C$, nadtlenek miedzi w temperaturach niższych od 4,1°C. Interesujące właściwości wykazuje żelazo w granicach temperatur 700–1000°C. Żelazo (odmiana α) ogrzewane do 768°C rozszerza się, następnie w granicach 768–906°C (odmiana β) kurczy się, a ogrzewane dalej znów się rozszerza (odmiana γ). Te trzy odmiany żelaza różnią się budową wewnętrzną i to decyduje o ich różnych właściwościach fizycznych i chemicznych. Na przykład tylko Fe_α ma właściwości magnetyczne.

14.2. Rozszerzalność objętościowa ciał stałych i ciekłych

Zjawisko cieplnej rozszerzalności objętościowej występuje we wszystkich ciałach, niezależnie od ich stanu skupienia. Zaczniemy od praw rozszerzalności objętościowej ciał stałych i ciekłych. Rozszerzalność objętościową ciała w granicach temperatur t_0 i t charakteryzuje wartość średniego współczynnika rozszerzalności objętościowej $\bar{\gamma}$:

$$\bar{\gamma} = \frac{V - V_0}{V_0 \Delta t},$$

gdzie V_0 oznacza objętość w temperaturze t_0, V — objętość w temperaturze t.

Ze wzoru na $\bar{\gamma}$ wynika, że

$$V = V_0(1 + \bar{\gamma}\Delta t). \tag{14.4}$$

Wartość średniego współczynnika $\bar{\gamma}$ wyraża, o jaką część objętości pierwotnej zwiększa się średnio objętość danego ciała przy wzroście temperatury o 1 kelwin w granicach temperatur od t_0 do t.

Możemy też określić współczynnik rozszerzalności w temperaturze t_0 jako granicę $\bar{\gamma}$, gdy Δt zmierza do zera:

$$\gamma = \lim_{\Delta t \to 0} \frac{V-V_0}{V_0 \Delta t}, \quad \text{skąd} \quad \gamma = \frac{1}{V_0}\frac{dV}{dt}. \tag{14.5}$$

W przypadku ciał stałych jednorodnych i izotropowych istnieje określona zależność między wartościami współczynników $\bar{\alpha}$ i $\bar{\gamma}$. Weźmy pod uwagę sześcian o krawędzi l_0 w temperaturze 0°C z materiału jednorodnego i izotropowego. Jego objętość początkowa

$$V_0 = l_0^3.$$

Po ogrzaniu do temperatury t, czyli o Δt, nowa długość krawędzi l będzie równa

$$l = l_0(1+\bar{\alpha}\Delta t),$$

a zatem objętość V po ogrzaniu

$$V = l_0^3(1+\bar{\alpha}\Delta t)^3.$$

Równocześnie obowiązuje zależność

$$V = V_0(1+\bar{\gamma}\Delta t).$$

Z porównania prawych stron obu równań wynika

$$V_0(1+\bar{\gamma}\Delta t) = l_0^3(1+3\bar{\alpha}\Delta t+3\bar{\alpha}^2\Delta t^2+\bar{\alpha}^3\Delta t^3).$$

Uwzględniając, że $V_0 = l_0^3$ oraz pomijając wyrazy z $\bar{\alpha}^2$ i $\bar{\alpha}^3$ jako bardzo małe wyższego rzędu otrzymujemy przybliżoną zależność

$$\bar{\gamma} \cong 3\bar{\alpha}.$$

W związku z tym w tabeli współczynników rozszerzalności objętościowej (tab. 14.2) nie umieściliśmy ciał stałych.

Tabela 14.2

Średnie współczynniki rozszerzalności objętościowej cieczy

Rodzaj cieczy	Granice temperatur °C	$\bar{\gamma}$ K^{-1}
tlen	−205−−184	0,00385
eter etylowy	0–30	0,00162
alkohol etylowy	0–30	0,00110
gliceryna	18	0,00053
rtęć	0–50	0,000182

14.3. Zmiany gęstości ciał podczas ogrzewania

Zmianom objętości ciał podczas ogrzewania towarzyszą zmiany gęstości. Aby je dokładniej opisać, weźmy pod uwagę ciało stałe lub ciekłe o masie m, mające w temperaturze t_0 objętość V_0 i gęstość ϱ_0. W temperaturze t jego objętość jest V. Nową gęstość oznaczamy przez ϱ.

Z definicji gęstości $\varrho = m/V$ wynika, że

$$V = \frac{m}{\varrho}, \quad \text{a} \quad V_0 = \frac{m}{\varrho_0}.$$

Podstawiając do wzoru (14.4) otrzymujemy

$$\frac{m}{\varrho} = \frac{m}{\varrho_0}(1+\bar{\gamma}\Delta t),$$

skąd

$$\varrho = \frac{\varrho_0}{1+\bar{\gamma}\Delta t}. \tag{14.6}$$

14.4. Właściwości cieplne gazów

Ogrzewanie gazów ma wpływ na ich objętość i prężność (czyli ciśnienie). Obie te zależności rozważymy oddzielnie, rozpatrując w odniesieniu do stałej masy gazu:

a) zmiany objętości podczas ogrzewania pod stałym ciśnieniem (przemiana izobaryczna),

b) zmiany prężności podczas ogrzewania w stałej objętości (przemiana izochoryczna).

Zmiany objętości gazów podczas ogrzewania pod stałym ciśnieniem można charakteryzować za pomocą średniego współczynnika rozszerzalności objętościowej, a zmiany prężności podczas ogrzewania w stałej objętości — za pomocą średniego współczynnika prężności $\bar{\beta}$. Średni współczynnik prężności $\bar{\beta}$ w granicach temperatur od t_0 do t wyraża się stosunkiem przyrostu prężności do prężności początkowej i przyrostu temperatury:

$$\bar{\beta} = \frac{p-p_0}{p_0\Delta t}. \tag{14.7}$$

Badania doświadczalne wykazały, że różne gazy silnie rozrzedzone, ogrzewane w tych samych granicach temperatur, mają prawie jednakowe średnie współczynniki rozszerzalności $\bar{\gamma}$ i prężności $\bar{\beta}$. Pamiętamy, że w odniesieniu do przemian izotermicznych przyjęto za obowiązujące prawo Boyle'a–Mariotte'a, przybliżone dla gazów rzeczywistych, dokładnie zaś stosujące się do gazu fikcyjnego — gazu doskonałego. Tę samą koncepcję gazu doskonałego wykorzystano przy ustalaniu praw rządzących przemianą izobaryczną i izochoryczną. Są to: *prawo Gay-Lussaca dla przemiany izobarycznej* —

$$V = \gamma V_0 T, \tag{14.8}$$

gdy $p = \text{const}$, *oraz prawo Charlesa dla przemiany izochorycznej* —

$$p = \beta p_0 T, \tag{14.9}$$

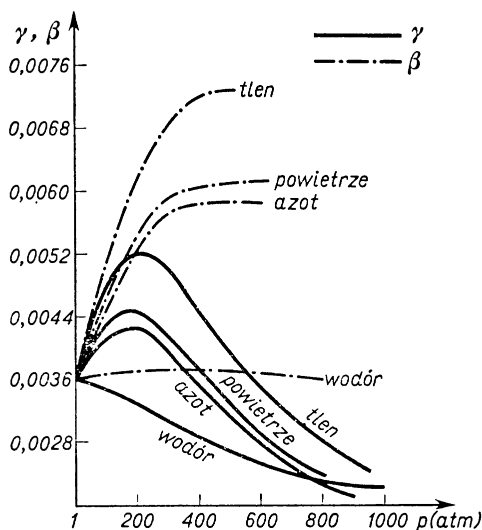
gdy $V = \text{const}$. Podobnie jak prawo Boyle'a–Mariotte'a, oba wymienione prawa stosują się do stałej masy gazu doskonałego*. Występujące w (14.8) i (14.9) współczynniki γ i β dla gazu doskonałego są jednakowe i mają wartość

$$\gamma = \beta \cong 0{,}003\,660\,4\ \text{K}^{-1} = \frac{1}{273{,}15}\text{K}^{-1},$$

* Do koncepcji gazu doskonałego wrócimy jeszcze w § 14.5 i 14.6 oraz 16.2.

gdy objętość początkowa V_0 i ciśnienie początkowe p_0 odniesione są do temperatury 0°C.

Prawa gazów doskonałych można traktować jako prawa graniczne, stosujące się do gazów rzeczywistych w wysokich temperaturach i przy ciśnieniach zmierzających do zera. Niektóre gazy rzeczywiste, a przede wszystkim wodór, hel i azot pod ciśnieniami nie przekraczającymi jednej atmosfery i w dość wysokich temperaturach wykazują małe odchylenia od właściwości gazu doskonałego. Dane z tab. 14.3 charakteryzują właściwości tych gazów przy ciśnieniach rzędu kilkudziesięciu centymetrów słupa rtęci i w temperaturach 0–100°C.



Rys. 14.1

Wpływ wysokiego ciśnienia w granicach temperatur ok. 273–373K na wartość współczynników rozszerzalności i prężności gazów przedstawia rys. 14.1. Wpływ temperatury na wartość omawianych współczynników charakteryzują następujące liczby: dla powietrza pod ciśnieniem 15 atm współczynnik γ zmienia się od 0,00427 K^{-1} w temperaturze —145°C do 0,00379 K^{-1} w temperaturze 100°C. Zmiana wynosi około 11%.

Tabela 14.3

Średnie współczynniki rozszerzalności $\bar{\gamma}$ i prężności $\bar{\beta}$ w K^{-1} gazów w granicach temperatur 273–373K

Ciśnienie cm Hg	Hel		Wodór		Azot	
	$\bar{\gamma}$	$\bar{\beta}$	$\bar{\gamma}$	$\bar{\beta}$	$\bar{\gamma}$	$\bar{\beta}$
39,024	0,0036595	0,0036617	0,0036604	0,0036617	0,0036664	0,0036673
53,311	36594	36602	36604	36613	36668	36671
72,717	36597	36611	36593	36620	36699	36709

14.5. Równanie stanu gazu doskonałego

W paragrafie 14.4 wymieniliśmy trzy prawa rządzące gazem doskonałym: prawo przemian izotermicznych, czyli prawo Boyle'a–Mariotte'a, prawo przemian izobarycznych, czyli prawo Gay-Lussaca oraz prawo przemian izochorycznych, czyli prawo Charles'a. Każde z wymienionych trzech praw ma narzucony pewien warunek ograniczający jego stosowalność, a mianowicie w każdej z wymienionych przemian jeden z trzech parametrów określających stan gazu, czyli jedna z wielkości T, V lub p jest stała.

Dążymy do znalezienia ogólnego prawa wiążącego temperaturę, ciśnienie i objętość danej masy gazu w stanie początkowym z wartościami ciśnienia, objętości i temperatury w stanie końcowym. Niech pewna masa gazu doskonałego zajmuje w temperaturze $T_0 = 273{,}15\text{K}$ objętość V_0 pod ciśnieniem p_0. W temperaturze wyższej T ta sama masa gazu zajmuje objętość V pod ciśnieniem p. Podczas tej przemiany zmieniły się wszystkie trzy parametry gazowe, nie można zatem stosować żadnego z wymienionych wyżej praw gazowych. W związku z tym wprowadźmy pomyślany stan przejściowy taki, aby jeden z parametrów w stanie pośrednim pozostał niezmieniony w stosunku do stanu początkowego, drugi osiągnął już wartość ostateczną, odpowiadającą stanowi końcowemu, a trzeci przyjął jakąś wartość przejściową.

Zestawmy charakterystyki wszystkich trzech stanów:

stan początkowy I

$$V_0 \quad p_0 \quad T_0 = 273{,}15\text{K}$$

stan przejściowy II

$$V_0' \quad p \quad T_0 = 273{,}15\text{K}$$

stan końcowy III

$$V \quad p \quad T$$

Przeanalizujmy kolejno procesy przejścia od stanu I do II, a następnie od stanu II do III.

Proces przejścia od stanu I do stanu II jest przemianą izotermiczną, a więc

$$V_0 p_0 = V_0' p,$$

skąd

$$V_0' = \frac{V_0 p_0}{p}.$$

Proces przejścia od stanu II do stanu III jest przemianą izobaryczną, a więc

$$V = \gamma V_0' T.$$

Po podstawieniu V_0' znajdujemy

$$V = \frac{V_0 p_0}{p} \gamma T,$$

ale

$$\gamma = \frac{1}{273{,}15} = \frac{1}{T_0},$$

a więc
$$\frac{Vp}{T} = \frac{V_0 p_0}{T_0}. \tag{14.10}$$

Jest to już równanie wiążące ze sobą trzy zmienne parametry, z tym jednak zastrzeżeniem, że temperatura gazu w stanie początkowym równa się $T_0 = 273{,}15\,\mathrm{K}$. Aby uwolnić się od tego warunku ograniczającego, badamy przemianę tej samej masy gazu od tego samego stanu początkowego

$$V_0 \quad p_0 \quad T_0 = 273{,}15\,\mathrm{K}$$

do nowego stanu końcowego

$$V_1 \quad p_1 \quad T_1$$

Rozumowanie analogiczne do poprzedniego doprowadzi nas do wniosku, że

$$\frac{V_1 p_1}{T_1} = \frac{V_0 p_0}{T_0}. \tag{14.11}$$

Z równości prawych stron równań (14.10) i (14.11) wynika, że

$$\frac{Vp}{T} = \frac{V_1 p_1}{T_1}. \tag{14.12}$$

Z równania tego, zwanego równaniem charakterystycznym lub *równaniem stanu gazu doskonałego*, wynika, że w dowolnej przemianie danej masy gazu doskonałego *stosunek iloczynu pV do temperatury w skali Kelvina pozostaje stały.*

Wracając do równania (14.10) znajdujemy

$$Vp = \frac{p_0 V_0}{273{,}15} T \quad \text{lub} \quad Vp = AT.$$

Wartość $A = p_0 V_0 / 273{,}15$ jest stała dla danej masy gazu. Nie jest to jednak stała uniwersalna, gdyż A zależy od masy i rodzaju badanego gazu.

Znając równanie stanu gazu można rozwiązywać różne zagadnienia. Wymienimy kilka przykładów.

1. Dość często nasuwa się konieczność redukcji objętości masy gazu do tzw. warunków normalnych, tzn. do ciśnienia 1 atm i temperatury 0°C (273,15K). Musimy wtedy znać wartości V, p i T, a szukamy V_0 pod ciśnieniem 1 atm i w temperaturze 0°C.

2. Innym zagadnieniem jest obliczanie masy badanego gazu o znanej gęstości ϱ_0 w warunkach normalnych. Pamiętając, że $\varrho_0 = m/V_0$ z równania

$$Vp = \frac{m p_0}{\varrho_0 \cdot 273{,}15} T$$

znajdujemy
$$m = \frac{\varrho_0 \cdot 273{,}15\, Vp}{p_0 T}.$$

3. Można też obliczać gęstość gazu ϱ w temperaturze t i pod ciśnieniem p, gdy znamy jego gęstość ϱ_0 w warunkach normalnych.

Obie strony równania

$$Vp = \frac{V_0 p_0}{273,15} T$$

dzielimy przez m. Stąd

$$\frac{1}{\varrho} p = \frac{1}{\varrho_0} \frac{p_0}{273,15} T,$$

czyli

$$\varrho = \frac{273,15 \varrho_0 p}{p_0 T}.$$

14.6. Równanie Clapeyrona

W równaniu

$$pV = AT$$

stała A jest zależna od masy i rodzaju badanego gazu. Zbadajmy to równanie w odniesieniu do jednego mola (jednej gramocząsteczki) dowolnego gazu. Z prawa Avogadra wiadomo, że w *jednakowych objętościach różnych gazów mierzonych w tych samych warunkach temperatury i ciśnienia znajduje się jednakowa liczba cząsteczek.* Poza tym wiemy, że liczba Avogadra $N = 6,02 \cdot 10^{23}$ określa liczbę cząsteczek zawartych w jednym molu dowolnego ciała. Z zestawienia tych wiadomości wynika, że jeden mol dowolnego gazu zajmuje w normalnych warunkach tę samą objętość.

Gazy rzeczywiste wykazują niewielkie odstępstwa od tej prawidłowości, jak to widać z tab. 14.4. Przyjmujemy, że objętość jednego mola gazu doskonałego w warunkach normalnych wynosi 0,022 414 m^3.

Tabela 14.4

Objętości molowe różnych gazów w warunkach normalnych

Rodzaj gazu	Masa molowa	Objętość molowa m^3
wodór	2,016	0,0224 20
azot	28,0	0,0224 10
tlen	32,0	0,0223 90
dwutlenek węgla	44,0	0,0222 60

W równaniu $pV = AT$ odniesionym do jednego mola stała A przekształca się na *stałą uniwersalną* R, zwaną *stałą gazową*:

$$R = \frac{p_0 V_{0m}}{273,15}, \tag{14.13}$$

gdzie p_0 oznacza ciśnienie 1 atm, a V_{0m} — objętość jednego mola gazu doskonałego w warunkach normalnych.

Po podstawieniu do (14.13) danych liczbowych i wymiarów na wartość stałej w jednostkach SI otrzymujemy

$$R = 8,314\,32 \text{ J/mol} \cdot \text{K}.$$

Należy pamiętać, że w równaniu

$$pV_m = RT, \tag{14.14}$$

zwanym równaniem Clapeyrona, V_m oznacza objętość molową danego gazu.

W praktyce rzadko mamy do czynienia z masami gazu dokładnie odpowiadającymi jednemu molowi. W odniesieniu do dowolnej masy m gramów równanie Clapeyrona przyjmuje inną postać. A mianowicie, jeżeli objętość V_m jednego mola, czyli M gramów gazu, wynosi

$$V_m = \frac{RT}{p},$$

to objętość V dowolnej masy m gramów

$$V = \frac{m}{M}\frac{RT}{p}.$$

Stąd

$$Vp = \frac{m}{M}RT. \tag{14.15}$$

Jest to równanie Clapeyrona odniesione do dowolnej masy m gramów gazu doskonałego, którego masa molowa wynosi M. Stosunek $m/M = n$ oznacza liczbę moli gazu w objętości V, a zatem

$$Vp = nRT. \tag{14.16}$$

Równanie Clapeyrona stosuje się do gazów rzeczywistych tylko z pewnym przybliżeniem. Wielu uczonych usiłowało ustalić równanie opisujące dokładniej zależność między wielkościami p, V i T w odniesieniu do gazów rzeczywistych. Jedna z takich bardziej udanych prób wiąże się z nazwiskiem van der Waalsa (por. § 17.12).

Na zakończenie jeszcze jedna uwaga. Mówiliśmy o trzech parametrach p, V, T, charakteryzujących stan takiego prostego układu, jakim jest określona masa gazu doskonałego. Obecnie — po ustaleniu równania stanu gazu doskonałego — łatwo stwierdzimy, że tylko dwa z nich są parametrami niezależnymi. Wartość trzeciego — przy danych dwóch — wynika jednoznacznie z równania stanu.

Pytania i zadania

1. Jakie jest znaczenie fizyczne współczynników rozszerzalności termicznej: a) liniowej, b) objętościowej?

2. Czy istnieje związek między współczynnikiem rozszerzalności liniowej i objętościowej izotropowego, jednorodnego ciała stałego lub ciekłego?

3. Jakie wzory wyrażają długość i objętość ciała po ogrzaniu?

4. Jakie jest znaczenie fizyczne współczynnika prężności gazów?

5. Jaką prawidłowość stwierdzono przy pomiarach współczynnika rozszerzalności i prężności gazów pod małymi ciśnieniami i w dość wysokich temperaturach?

6. Jakie rodzaje przemian gazowych można rozróżniać zakładając stałość jednego z parametrów określających stan gazu? Jakie prawa stosują się do tych przemian?

7. Jakie rozumowanie doprowadza do równania stanu gazu doskonałego? (Podać plan wyprowadzenia wzoru.)

8. Na czym polega przejście od równania stanu gazu doskonałego, odniesionego do dowolnej masy gazu, do równania Clapeyrona?

9. Co oznacza i jaką wartość liczbową w układzie SI ma stała gazowa?

10. Odległość dwóch punktów mierzona stalową taśmą w temperaturze 30°C wynosi 200 m. Jaka jest rzeczywista odległość tych punktów, jeśli taśma była cechowana w temperaturze 20°C. Współczynnik rozszerzalności liniowej stali $\alpha = 12 \cdot 10^{-6}$ K^{-1}.

Odp. 200,024 m.

11. Stalowy pręt o przekroju $S = 2$ cm^2 ogrzano od 0 do 50°C i następnie gwałtownie oziębiono do 0°C. Jaką najmniejszą siłą należy działać wzdłuż osi pręta, aby w czasie oziębiania nie uległ on skróceniu? Zakładamy, że moduł Younga nie zależy od temperatury w badanym zakresie i równa się $E = 2,1 \cdot 10^{11}$ N/m^2. Współczynnik rozszerzalności liniowej stali $\alpha = 12 \cdot 10^{-6}$ K^{-1}.

Odp. 25200 N.

12. Wahadło mosiężne w postaci pręta o długości l jest w temperaturze 15°C wahadłem sekundowym (tj. czas jego wahnięcia $T/2 = 1$ s). Obliczyć okres tego wahadła w temperaturze 25°C. Moment bezwładności tego wahadła względem osi obrotu $I = ml^2/3$, a odległość środka masy od osi obrotu $a = l/2$. Współczynnik rozszerzalności liniowej mosiądzu $\alpha = 19 \cdot 10^{-6}$ K^{-1}.

Odp. 2,00008 s.

13. Wiedząc, że gęstość rtęci w temperaturze 0°C wynosi $\varrho = 13,59$ g/cm^3, obliczyć jej gęstość w temperaturze 10 i 100°C. Współczynnik rozszerzalności objętościowej rtęci $\gamma = 18,2 \cdot 10^{-5}$ K^{-1}.

Odp. 13,57 g/cm^3; 13,34 g/cm^3.

14. Kulkę szklaną o współczynniku rozszerzalności objętościowej γ_s zanurzamy dwukrotnie w tej samej cieczy: raz w temperaturze t, drugi raz w temperaturze t_1. Wypory w obu przypadkach wynoszą odpowiednio P i P_1. Wyznaczyć współczynnik rozszerzalności objętościowej cieczy.

Odp. $\dfrac{P-P_1+P\gamma_s(t_1-t)}{P_1(t_1-t)}$.

15. O ile należy zwiększyć ciśnienie zewnętrzne, aby podczas ogrzewania rtęci od 10 do 20°C objętość jej nie uległa zmianie. Współczynnik rozszerzalności objętościowej rtęci $\gamma = 18,2 \cdot 10^{-5}$ K^{-1}; moduł ściśliwości rtęci $= 39 \cdot 10^{-6}$ atm^{-1}.

Odp. 46,6 atm.

16. Jaka jest gęstość powietrza w temperaturze $t_1 = 27$°C i pod ciśnieniem $p_1 = 5$ atm, jeśli jego gęstość w warunkach normalnych $\varrho_0 = 0,001\,293$ g/cm^3.

Odp. 0,005 88 g/cm^3.

17. W butli stalowej znajduje się gaz o temperaturze 27°C pod ciśnieniem 40 atm. Jak się zmieni ciśnienie gazu, jeśli z butli ulotni się połowa zawartej w niej pierwotnie masy gazu i temperatura obniży się o 15°C.

Odp. Zmniejszy się o 21 atm.

18. Manometr dołączony do butli ze sprężonym gazem wskazuje w temperaturze $t_1 = 27$°C ciśnienie $p_1 = 100 \cdot 10^5$ N/m^2. Jakie ciśnienie będzie on wskazywał, gdy temperatura obniży się do $t_2 = -23$°C. Zmiany objętości butli zaniedbać.

Odp. 83,3 $\cdot$ 10^5 N/m^2.

19. Balon o pojemności 20 l napełniony jest sprężonym powietrzem. Dołączony do balonu manometr w temperaturze 20°C wskazywał 120 $\cdot$ 10^5 N/m^2. Jaką objętość wody można usunąć za pomocą tego powietrza z cysterny łodzi podwodnej, jeśli łódź znajduje się na głębokości 30 m i temperatura balonu wynosi 5°C. Przyjąć, że ciśnienie słupa wody o wysokości 10 m równa się 10^5 N/m^2.

Odp. ok. 550 litrów.

ROZDZIAŁ 15

Kalorymetria

15.1. Jednostka ilości ciepła i ciepło właściwe

Tytuł niniejszego rozdziału nawiązuje do nazwy dawnej jednostki ilości ciepła — kalorii. Po wprowadzeniu układu SI odrębna jednostka ilości ciepła wychodzi z użycia, gdyż ciepło jako jeden z rodzajów energii może być wyrażane w jednostkach energii (pracy), tzn. w dżulach i jego pochodnych*.

W rozdziale poświęconym termometrii była mowa o prawie wyrównywania się temperatur. Wyrównywanie się temperatur jest możliwe dzięki przekazywaniu energii cieplnej od jednego ciała czy układu ciał do drugiego. Możliwość ścisłego określania ilości ciepła oddawanego lub pobieranego opiera się na następujących podstawach:

1. Ciepło pobrane przez ciało A podczas ogrzania od temperatury T_1 do temperatury T_2 równa się ciepłu oddanemu przez to samo ciało A podczas oziębiania od temperatury T_2 do T_1, jeśli w obu przypadkach ciało przechodzi przez te same stany przejściowe.

2. Jeżeli ciało A kilkakrotnie w różnych okolicznościach, np. podczas zetknięcia kolejno z różnymi ciałami B, C, D, ..., ogrzewa się od temperatury T_1 do T_2 przechodząc przez te same stany pośrednie, to za każdym razem, niezależnie od rodzaju ciała udzielającego ciepła, pobierana jest taka sama ilość ciepła.

3. *Ilość ciepła ΔQ (lub po prostu Q) pobierana przez ciało jednorodne przy niewielkim wzroście temperatury, przy którym nie zachodzi przemiana fazowa, jest proporcjonalna do masy ciała i do przyrostu temperatury* $\Delta T = T_2 - T_1$. Współczynnikiem proporcjonalności jest ciepło właściwe średnie $\bar{c}$, odpowiadające przedziałowi temperatur T_1 i T_2:

$$\Delta Q = m\bar{c}\Delta T. \tag{15.1}$$

Stąd *ciepło właściwe średnie* w granicach temperatur T_1 i T_2

$$\bar{c} = \frac{\Delta Q}{m\Delta T}.$$

* Ponieważ w wielu pracach ilości ciepła są podawane jeszcze w kaloriach i kilokaloriach, więc przypominamy, że

1 cal = 4,1868 J,
1 kcal = 4186,8 J.

Zmniejszając nieskończenie wartość przyrostu temperatury znajdujemy ciepło właściwe ciała w temperaturze T jako granicę wartości $\bar{c}$:

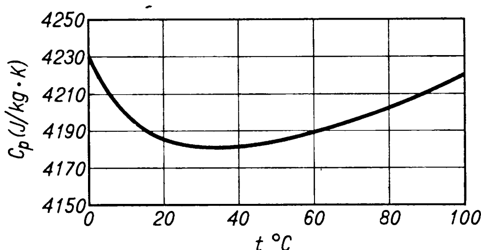
$$c = \lim_{\Delta T \to 0} \frac{\Delta Q}{m \Delta T} = \frac{1}{m} \frac{dQ}{dT}. \tag{15.2}$$

Z powyższych równań wynika, że *ciepło właściwe* dowolnego ciała wyraża liczbowo ilość ciepła potrzebną do ogrzania 1 kg ciała o 1 K. Jednostką ciepła właściwego jest J/kg · K.

Rozpatrzmy bliżej ciepło właściwe wody. Jest to wielkość zależna od temperatury. Ogrzanie 1 kg wody destylowanej od 14,5 do 15,5°C (czyli od 287,6–288,6K) wymaga doprowadzenia energii cieplnej w ilości 4186,8 J. W przeliczeniach nie wymagających wielkiej dokładności można przyjmować, że ciepło właściwe wody jest stałe w granicach temperatury 0–100°C. Dokładne badania wykazują, że przy takim założeniu można popełnić błąd do 1%. Rys. 15.1 przedstawia charakter zależności ciepła właściwego wody c_p od temperatury t. Zależność ciepła właściwego od temperatury dla większości ciał w temperaturach bliskich 0°C można wyrazić następującym równaniem:

$$c = \alpha + \beta t + \gamma t^2,$$

gdzie α, β, γ są stałe dla danego ciała.



Rys. 15.1

W temperaturach bliskich zera bezwzględnego ciepło właściwe może być wyrażo *wzorem Debye'a*:

$$c = aT^3,$$

gdzie a oznacza współczynnik proporcjonalności zależny od rodzaju ciała, T — temperaturę w skali Kelvina.

Ciepło właściwe jest nie tylko funkcją temperatury. Zależy ono także od tego, czy ogrzewanie ciała odbywa się z zachowaniem stałej objętości, czy też pod stałym ciśnieniem. Podczas ogrzewania ciała z zachowaniem stałej objętości ciepło musi być dostarczane jedynie na wywołanie przyrostu energii wewnętrznej związanego ze wzrostem temperatury, natomiast podczas ogrzewania pod stałym ciśnieniem ciało rozszerza się i pokonuje opór wynikający z istnienia zewnętrznego ciśnienia, a zatem wykonuje pracę. W tym przypadku dostarczona energia cieplna zużywa się częściowo na ogrzanie ciała (tzn. na zwiększenie energii wewnętrznej, podobnie jak przy ogrzewaniu

w stałej objętości), a częściowo na wykonanie pracy związanej ze zmianą objętości. *Ciepło właściwe pod stałym ciśnieniem c_p jest zatem większe od ciepła właściwego w stałej objętości c_V:*

$$c_p > c_V.$$

W przypadku ciał stałych i ciekłych zmiany objętości podczas ogrzewania są niewielkie, a więc różnice c_p i c_V są nieznaczne i czasem w ogóle nie są brane pod uwagę. W tabelach np. często są podawane pojedyncze wartości ciepła właściwego dla ciał stałych i ciekłych, w zasadzie odpowiadające wartości c_p (tab. 15.1).

Tabela 15.1

Ciepło właściwe

Rodzaj substancji	Temperatura °C	c_p J/kg · K	c_V J/kg · K
a) ciała stałe i ciekłe			
cynk	20	384	
glin	20	897	
miedź	20	381	
mosiądz	20	387	
ołów	20	125	
stal	20	461	
platyna	500	1470	
lód	−20–0	2095	
szkło	20	837	
woda	0–100	4186	
alkohol etylowy	0	2240	
rtęć	20	1390	
b) ciała gazowe			
azot	15	1043	742
tlen	15	913	649
powietrze	15	1006	729
hel	−180	5240	3155
wodór	20	14250	10050

Inaczej przedstawia się sprawa w przypadku gazów, gdzie wyraźnie należy rozróżniać c_p i c_V.

15.2. Pomiar ilości ciepła

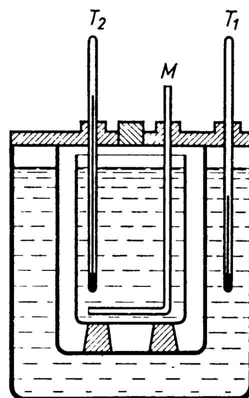
Pomiary ilości ciepła powinny być wykonywane w *układach doskonale izolowanych cieplnie od otoczenia*, tzn. idealnie zabezpieczonych przed dopływem ciepła z zewnątrz lub odpływem ciepła na zewnątrz. Wtedy można wewnątrz układu wyodrębnić ciała o temperaturze wyższej — oddające ciepło i ciała o temperaturze niższej — pobierające ciepło. W układzie izolowanym cieplnie obowiązuje *zasada zachowania ilości ciepła*. Odbywa się tylko wymiana energii cieplnej między ciałami o różnych temperaturach,

ale całkowite ciepło pobrane przez ciała o temperaturze niższej równa się całkowitemu ciepłu oddanemu przez ciała o temperaturze wyższej. Jest to tzw. *zasada bilansu cieplnego*, stanowiąca odpowiednik zasady zachowania energii w odniesieniu do zjawisk cieplnych.

Jeżeli żadne z ciał biorących udział w wymianie ciepła nie ulega przemianie fazowej (por. § 17.1 i 17.4), *to każdy ze składników bilansu cieplnego ma postać trójczynnikowego iloczynu, a mianowicie*

$$\Delta Q = cm\Delta T.$$

Urządzeniami zapewniającymi dobrą izolację cieplną są kalorymetry różnych typów, stosowane szeroko do pomiarów ilości ciepła. Najbardziej rozpowszechniony jest kalorymetr wodny (rys. 15.2). Jest to podwójne naczynie metalowe o jasnych, wypolerowanych ścianach. Naczynie wewnętrzne, zwane czasem kalorymetrem właściwym, ustawione jest na podstawkach z dobrego izolatora cieplnego, np. z korka. Warstwa powietrza między ściankami obu naczyń również zabezpiecza przed stratami ciepła. Całość urządzenia czasem jest umieszczana w trzecim naczyniu wypełnionym wodą i jest przykryta drewnianą pokrywą z otworami na mieszadełko i termometr.



Rys. 15.2

Opiszemy pomiar ciepła właściwego ciała stałego wykonany za pomocą kalorymetru wodnego. Wyznaczamy masę kalorymetru właściwego łącznie z mieszadełkiem, m_k, następnie masę kalorymetru z mieszadełkiem i wodą. Stąd znajdujemy masę m_w wody wlanej do kalorymetru. Na termometrze zaopatrzonym w podziałkę o dokładności do 0,1 stopnia Kelvina (lub Celsjusza) wyznaczamy temperaturę początkową wody i kalorymetru T_0 (lub t_0). Następnie wrzucamy do kalorymetru ciało o masie m, ogrzane np. w kąpieli wodnej do określonej temperatury T (lub t). Po parominutowym mieszaniu, które przyspiesza wymianę ciepła, ustala się nowa równowaga termiczna i odczytujemy końcową temperaturę T_k (lub t_k).

Zakładając, że ciałami biorącymi udział w wymianie ciepła są tylko kalorymetr, mieszadełko, termometr, woda i badane ciało stałe, układamy bilans cieplny. Ilość ciepła ΔQ_1 oddana przez ciało badane równa się ilości ciepła ΔQ_2 pobranej przez kalorymetr, mieszadełko, termometr i wodę:

$$\Delta Q_1 = mc_x(T-T_k)^*,$$

gdzie c_x oznacza szukane ciepło właściwe.

$$\Delta Q_2 = m_k c_k(T_k-T_0)+m_w c_w(T_k-T_0)+m_t c_t(T_k-T_0),$$

$$mc_x(T-T_k) = (m_k c_k+m_w c_w+m_t c_t)(T_k-T_0).$$

Stąd

$$c_x = \frac{(m_k c_k+m_w c_w+m_t c_t)(T_k-T_0)}{m(T-T_k)}.$$

W kalorymetrii często posługujemy się pojęciem pojemności cieplnej. Przez *pojemność cieplną* rozumiemy iloczyn mc, tzn. iloczyn masy ciała przez jego ciepło właściwe. Iloczyn ten równy dQ/dT (por. wzór (15.2)) wyraża liczbowo ilość ciepła potrzebną do ogrzania danego ciała o 1 K. Wyraża się w J/K. (Jeśli np. podana jest pojemność cieplna termometru, to iloczyn tej pojemności cieplnej przez przyrost temperatury wyraża ilość ciepła pobraną przez termometr.)

Pomiar ciepła właściwego gazów pod stałym ciśnieniem zazwyczaj jest wykonywany

* Przypominamy, że $\Delta T = \Delta t$.

następująco: z dużego zbiornika z regulowanym zaworem odpływa do kalorymetru ogrzany gaz w ten sposób, że manometr wskazuje stałe ciśnienie. Odpowiednio umieszczony termometr mierzy temperaturę początkową gazu. W kalorymetrze gaz przechodzi przez metalową wężownicę zanurzoną w wodzie, oddając przy tym ciepło. Termometr umieszczony przy wylocie wężownicy pozwala wyznaczyć temperaturę końcową gazu. Znając masę gazu przepuszczonego przez wężownicę, jego temperaturę początkową i końcową oraz masę wody, kalorymetru i wężownicy, jak również ich temperatury na początku i przy końcu pomiaru, układamy bilans cieplny. Z równania bilansu znajdujemy ciepło właściwe gazu pod stałym ciśnieniem.

Badania licznych uczonych, a przede wszystkim fizyka francuskiego Regnaulta, wykazały, że c_p gazów mało zależy od ciśnienia w temperaturach dalekich od temperatury krytycznej (por. § 17.9). Gazy, takie jak wodór, hel, azot (zbliżone we właściwościach do gazu doskonałego), mają ciepło właściwe pod stałym ciśnieniem stosunkowo mało zależne także i od temperatury.

Bezpośredni pomiar ciepła właściwego gazów w stałej objętości jest mało dokładny. Stosując zbiorniki o małych objętościach mamy małe masy gazu (stąd duży błąd względny przy jej wyznaczaniu), a ze zbiornikami o dużych objętościach związana jest duża pojemność cieplna, co zwiększa błąd względny przy określaniu ilości ciepła pobranej przez masę gazu. W związku z tym wartość c_V wyznacza się zwykle pośrednio na podstawie pomiaru innych wielkości.

15.3. Prawo Dulonga i Petita

Do ciał stałych stosuje się przybliżone *prawo Dulonga i Petita*, według którego *iloczyn ciepła właściwego ciała stałego i masy jego gramoatomu jest wielkością stałą, równą* ok. 26 J/gramoatom · K, zwaną *ciepłem atomowym* i oznaczaną symbolem c_a. Ciepło atomowe jest to zatem pojemność cieplna jednego gramoatomu danej substancji. Tak np. w temperaturach bliskich pokojowej ciepła właściwe glinu i miedzi wynoszą odpowiednio 922 J/kg · K i 390 J/kg · K, masy zaś ich gramoatomów — 27 g/gramoatom i 63,5 g/gramoatom. Ich ciepła atomowe są więc odpowiednio równe 24,9 oraz 25,3 J/gramoatom · K.

Dość duże odstępstwa od prawa Dulonga i Petita obserwuje się w temperaturze pokojowej w przypadku pierwiastków lekkich, jak np. węgla i boru. W miarę wzrostu temperatury ciepło atomowe nawet i tych pierwiastków zbliża się do wartości 26 J/gramoatom · K*.

W pobliżu zera bezwzględnego ciepło atomowe wszystkich ciał maleje zbliżając się do zera w sposób mniej lub bardziej gwałtowny.

Wiadomo że gramoatomy dowolnych pierwiastków składają się z tej samej liczby atomów. Z prawa Dulonga i Petita wypływa zatem wniosek, że *do ogrzania takiej samej liczby atomów o taką samą liczbę stopni potrzebne są jednakowe ilości ciepła, niezależnie od natury chemicznej atomów.*

---

* Wymieniona wartość ciepła atomowego odpowiada 6,3 cal/gramoatom · K.

Pytania i zadania

1. Dlaczego ciepło można wyrażać w jednostkach pracy?

2. Co nazywamy ciepłem właściwym danego ciała? Co oznaczają symbole c_p i c_V? W odniesieniu do jakich ciał jest specjalnie ważne odróżnianie ciepła właściwego c_p i c_V?

3. Wyjaśnić dlaczego ciepło właściwe pod stałym ciśnieniem jest zawsze większe od ciepła właściwego przy stałej objętości.

4. Na czym się opieramy układając bilans cieplny?

5. Jak się nazywają przyrządy służące do pomiarów ilości ciepła? Podać przykłady, opisać budowę.

6. Opisać pomiar ciepła właściwego cieczy w kalorymetrze wodnym.

7. Co nazywamy pojemnością cieplną danego ciała? Jak się nazywa pojemność cieplna jednego gramoatomu dowolnego pierwiastka? Jakie prawo dotyczy tej pojemności cieplnej? Jaki jest głębszy sens fizyczny tego prawa?

8. Wyjaśnić dlaczego w przypadku gazów bezpośrednio wyznacza się raczej c_p niż c_V.

9. Do pieca hutniczego włożono platynową kulkę o masie 100 g. Natychmiast po wyjęciu jej z pieca wrzucono ją do mosiężnego kalorymetru o masie 250 g, zawierającego 1 kg wody o temperaturze 20°C. aka była temperatura pieca, jeżeli temperatura końcowa wody w kalorymetrze wynosiła 24°C? Potrzebne do rozwiązania zadania wartości ciepła właściwego podane są w tab. 15.1.

Odp. 1303°C.

10. Balon wypełniony tlenem pod ciśnieniem 80 atm w temperaturze 7°C został następnie ogrzany do temperatury 15,5°C. Przyjmując, że pojemność balonu jest stała i równa 10 l, obliczyć ilość ciepła pobraną przez tlen. Gęstość tlenu w warunkach normalnych wynosi 1,4 g/l; $c_V = 0{,}65$ J/g · K.

Odp. ok. 6000 J.

11. 400 g azotu ogrzano od 20 do 100°C pod stałym ciśnieniem. Jaka ilość ciepła została przy tym pochłonięta? Jaki był przyrost energii wewnętrznej gazu? Ile energii zużyło się na wykonanie pracy zewnętrznej związanej ze wzrostem objętości? Dane dla azotu: $c_p = 1{,}043$ J/g · K; $c_V = 0{,}742$ J/g · K.

Odp. 33 400 J, 23 750 J, 9650 J.

ROZDZIAŁ 16

Fizyka molekularna

16.1. Ogólne założenia teorii kinetyczno-molekularnej

Teorii kinetyczno-molekularnej zawdzięczamy specjalne, „mikroskopowe" ujęcie zjawisk zachodzących w otaczającym nas świecie. Teoria ta opiera się na następujących założeniach ogólnych:

1. Ciała mają budowę nieciągłą; składają się z drobnych elementów w postaci atomów lub cząsteczek (molekuł).

2. Wymienione elementy budowy ciał są w ciągłym ruchu. Wartości liczbowe i kierunki prędkości poszczególnych elementów są różne.

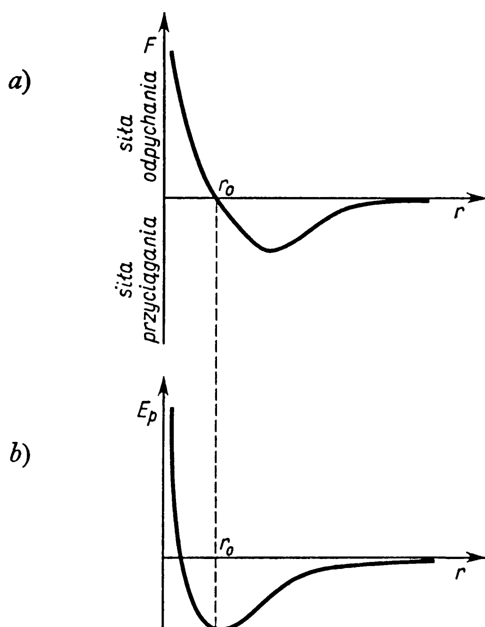
3. Pomiędzy poszczególnymi elementami budowy ciał występują siły wzajemnego oddziaływania.

Założenie pierwsze, dotyczące atomowej budowy materii, sięga swymi pierwszymi koncepcjami aż do czasów starożytnych (Demokryt, Epikur, Lukrecjusz), były to jednak w owym czasie tylko pewne spekulacje filozoficzne, nie oparte na danych doświadczalnych. Założenie to zostaje ponownie wysunięte w w. XVII i następnie, dzięki pracom licznych badaczy (Boyle, Bernoulli, Łomonosow, Dalton, Maxwell i wielu innych), przekształca się w poważną hipotezę naukową.

Założenie drugie znajduje potwierdzenie w licznych faktach doświadczalnych, jak np. w zjawisku rozprężania się gazów, w zjawisku dyfuzji (por. p. 16.11.2), w ruchach Browna (§ 16.15) itp.

Założenie trzecie dotyczy oddziaływań międzycząsteczkowych, które w dużym stopniu zależą od wzajemnych odległości między cząsteczkami i gwałtownie maleją ze wzrostem tych odległości. O charakterze tych sił będziemy jeszcze mówili. W tej chwili poprzestaniemy na stwierdzeniu, że właśnie różnice wartości sił międzycząsteczkowych, wynikające z różnego rzędu wielkości odległości międzycząsteczkowych w gazach z jednej strony, a w cieczach i ciałach stałych z drugiej, zdecydowały o tym, że teoria kinetyczna została najbardziej rozwinięta w odniesieniu do ciał gazowych. Z punktu widzenia odległości międzycząsteczkowych ciecze i ciała stałe traktujemy jako jedną grupę opierając się na tym, że zmiany objętości podczas procesu topnienia lub krzepnięcia są stosunkowo niewielkie.

Na przykładzie wody wyznaczymy średnią odległość między cząsteczkami a tym samym rząd wielkości odległości międzycząsteczkowych w ciałach stałych i ciekłych. Jeden mol wody, czyli około 18 g, zajmuje w temperaturze 4°C objętość ok. 18 cm^3 i zawiera około $6 \cdot 10^{23}$ cząsteczek. Na pojedynczą cząsteczkę przypada w tych warunkach objętość około $3 \cdot 10^{-29}$ m^3. Zakładając, że każda cząsteczka (traktowana jako punkt materialny) znajduje się w środku małego sześcianu, można obliczyć, że krawędź takiego sześcianu, a równocześnie odległość dwóch sąsiednich cząsteczek, równa się około $3 \cdot 10^{-10}$ m. Jest to też rząd wielkości odległości międzycząsteczkowych ciał stałych i ciekłych.



Rys. 16.1

Rysunki 16.1a i b ilustrują typowe przebiegi zależności sił międzycząsteczkowych i energii potencjalnej od wzajemnej odległości cząsteczek. (Wykres energii potencjalnej można otrzymać z wykresu siły pamiętając o tym, że siła $F = dE_p/dr$.) Z wykresu siły widać, że dla pewnej odległości $r = r_0$ siła staje się równa zeru — nie ma wzajemnego oddziaływania cząsteczek; tej odległości międzycząsteczkowej odpowiada stan równowagi trwałej układu. Z wykresu 16.1b widać, że tej odległości odpowiada minimum energii potencjalnej, co jest właśnie warunkiem trwałości równowagi.

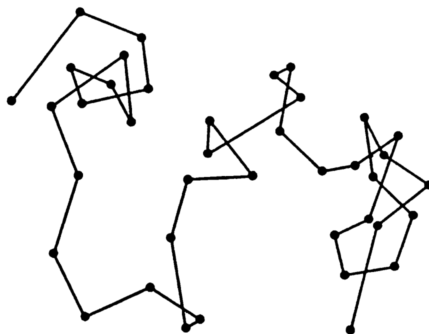
Odległościom $r < r_0$ odpowiadają siły odpychania gwałtownie rosnące przy malejącym r. Tym stanom układu odpowiadają duże wartości E_p. Dla $r > r_0$ i stopniowo rosnących charakterystyczne są siły przyciągania międzycząsteczkowego na razie rosnące, a potem dość szybko malejące do zera. Równocześnie E_p stopniowo od swej wartości minimalnej rośnie do zera.

Siły międzycząsteczkowe to typowe siły tzw. krótkiego zasięgu: duże — przy odległościach mniejszych od około 10^{-10} m, znikające praktycznie — przy odległościach rzędu 10^{-9} m.

Wyznaczona w naszym przykładzie odległość między cząsteczkami wody przypada na ten zakres odległości, któremu odpowiadają duże siły międzycząsteczkowe.

16.2. Wprowadzenie do teorii kinetycznej gazów

Ujęcie właściwości gazów z punktu widzenia teorii kinetycznej przedstawia się najprościej z tego względu, że dzięki stosunkowo dużym średnim odległościom międzycząsteczkowym w gazie można w pierwszym przybliżeniu zaniedbywać siły wzajemnego oddziaływania cząsteczek. Wyznaczmy podobnie jak poprzednio rząd wielkości średniej odległości międzycząsteczkowej w gazie. Jak wiadomo, jeden mol dowolnego gazu zajmuje w warunkach normalnych objętość $22{,}4 \cdot 10^{-3}$ m^3 i zawiera około $6 \cdot 10^{23}$ cząsteczek. A zatem średnia odległość międzycząsteczkowa w gazie w warunkach normalnych wynosi około $3 \cdot 10^{-9}$ m, czyli jest o jeden rząd wielkości większa niż w przypadku ciał stałych i ciekłych. Przy takich odległościach oddziaływanie wzajemne cząsteczek jest znikomo małe i można je zaniedbywać. Właśnie dzięki temu w dalszych rozważaniach będziemy zakładali, że każda cząsteczka gazu porusza się swobodnie (bez działania sił) aż do momentu zderzenia z inną cząsteczką albo ze ścianką naczynia. Stąd wniosek, że odcinki dróg między kolejnymi zderzeniami są przebywane ruchem jednostajnym prostoliniowym. Wobec zupełnej przypadkowości zderzeń drogi te mają różne długości. Obraz ruchu cząsteczki gazowej przedstawia rys. 16.2. Przyjmując upraszczające założenia, że: 1) *cząsteczki gazu wzajemnie na siebie nie działają aż do momentu zderzenia* i 2) *rozmiary cząsteczek można pominąć*, traktując je jako punkty (czyli zaniedbując objętość własną cząsteczek), ograniczamy nasze rozważania do *gazu doskonałego*.



Rys. 16.2

Cząsteczki podczas zderzeń zmieniają swe prędkości. Zderzenia cząsteczek gazowych można traktować jako zderzenia doskonale sprężyste (por. p. 8.4.1). Średnia długość prostoliniowego przebiegu między dwoma zderzeniami, wyliczona z bardzo wielkiej

liczby tych przelotów, nosi nazwę *średniej drogi swobodnej*. Jej wartość dla tlenu w warunkach normalnych wynosi $6{,}4 \cdot 10^{-8}$ m. W specjalnych badaniach wyznaczono też średnie prędkości ruchu cząsteczek. Tak np. średnia prędkość cząsteczki tlenu w temperaturze 273,15K (czyli 0°C) wynosi 425 m/s.

Znając średnią prędkość cząsteczki tlenu $\bar{v}$ oraz średnią drogę swobodną l można obliczyć średnią liczbę zderzeń $\bar{n}$ w czasie jednej sekundy:

$$\bar{n} = \frac{\bar{v}}{l} = \frac{425 \text{ m/s}}{6{,}4 \cdot 10^{-8} \text{ m}} = 6{,}6 \cdot 10^{9} \text{ s}^{-1}.$$

Przytoczone wartości liczbowe dla tlenu pozwalają wyciągnąć pewne ogólne wnioski co do wielkości charakteryzujących cząsteczki gazowe.

Liczba cząsteczek zawartych nawet w niewielkich objętościach rzędu 1 cm^3 w warunkach normalnych jest olbrzymia — rzędu 10^{19}. Liczba zderzeń przypadających na sekundę jest również bardzo duża, odpowiada miliardom. Te wartości odbiegają znacznie od liczb, z jakimi mamy do czynienia w życiu codziennym. Nic więc dziwnego, że teoria kinetyczna wykorzystuje specjalne metody rachunkowe, a mianowicie metody statystyczne, rachunek prawdopodobieństwa, operowanie wartościami średnimi itp. Aby lepiej sobie uzmysłowić znaczenie liczby rzędu 10^{19}, pójdźmy śladami Crookes'a i rozważmy taki przypadek. Mała ampułka szklana o objętości 1 cm^3 jest całkowicie opróżniona. Za pomocą np. iskry elektrycznej robi się w ściance ampułki otwór bardzo małych rozmiarów. Niech gaz przenika przez ten otwór do wnętrza ampułki w sposób jednostajny, np. w liczbie 100 milionów cząsteczek na sekundę. Proste przeliczenie wykazuje, że wewnątrz ampułki znajdzie się $2{,}7 \cdot 10^{19}$ cząsteczek (warunki normalne) po upływie $2{,}7 \cdot 10^{11}$ sekund, czyli po upływie około 9 tysięcy lat.

W tym przykładzie liczba cząsteczek wpływających do ampułki w każdej sekundzie była już bardzo duża, równa 100 milionom. Czy mamy „wyczucie" liczby 100 milionów? Na przykład, gdyby rozmiary podłużne komara ($\sim$1 cm) wzrosły 100 milionów razy, czy otrzymalibyśmy stworzenie o rozmiarach pytona, czy długiego pociągu towarowego? Najlepiej to sobie przeliczyć.

16.3. Rozkład prędkości Maxwella

Wobec olbrzymiego „tłoku" cząsteczek gazowych ruch ich, zwany często *ruchem cieplnym*, jest pod względem kierunku doskonale bezładny, chaotyczny. Należy to rozumieć w ten sposób, że nie ma w gazie *żadnego wyróżnionego kierunku prędkości. Wszystkie kierunki prędkości są równie prawdopodobne.* Gdyby nawet w pewnej chwili wystąpiło uprzywilejowanie pewnego kierunku prędkości, to olbrzymia liczba przypadkowych zderzeń międzycząsteczkowych niezwłocznie usunęłaby to uporządkowanie.

Inaczej przedstawia się sprawa występowania różnych wartości liczbowych prędkości. Pod tym względem nie mamy doskonałego chaosu, nie mamy zupełnego równouprawnienia wszystkich wartości liczbowych prędkości. Fizykowi angielskiemu Maxwellowi

zawdzięczamy rozwiązanie *zagadnienia dotyczącego prawdopodobieństwa występowania cząsteczek o różnych wartościach liczbowych prędkości.*

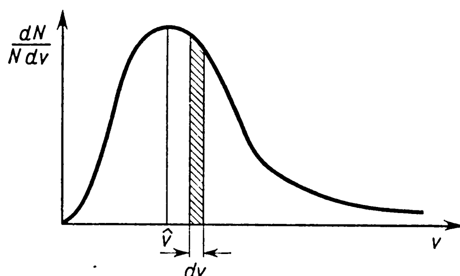
Prawo rozkładu prędkości Maxwella (uzupełnione następnie przez Boltzmanna) zostało wyprowadzone w oparciu o następujące założenia:

1. *Zderzenia cząsteczek mają charakter doskonale sprężysty.*

2. *Zgodnie z wymaganiami statystyki liczba cząsteczek dN, mających prędkości zawarte w dowolnie małym przedziale prędkości od v do $v+dv$, jest stała.*

Aby lepiej zrozumieć ten drugi warunek, rozważmy takie okoliczności. Niech liczba mieszkańców pewnej części świata wynosi $N = 300$ milionów. W tej liczbie występują ludzie rozmaitego wieku: od niemowląt do poważnych starców. Na podstawie danych biur ewidencji ludności ustalono liczbę dN ludzi, których wiek waha się w granicach 30–31 lat. Okazuje się, że obliczenia powtórzone za rok czy za dwa, czy za dziesięć, dają tę samą liczbę dN ludzi w wieku od 30 do 31. Nie będą to już ci sami ludzie, co w chwili przeprowadzania pierwszych obliczeń, bo ci już „się postarzeli", przeszli do grupy o innych granicach wieku. Na ich miejsce przyszli nowi z grupy 29, 28, czy 20-letnich. Ludzie objęci wskazanymi granicami wieku będą się starzeli, ale ogólna liczba objętych danymi granicami wieku będzie pozostawała niezmieniona. Wyniki będą tym dokładniejsze, im większą liczbę osób obejmie przeprowadzona statystyka.

W odniesieniu do wielkiej liczby cząsteczek gazowych sprawa przedstawia się podobnie. Niech dN oznacza liczbę cząsteczek o prędkościach zawartych w granicach v i $v+dv$. Na skutek nieuniknionych zderzeń pewna liczba cząsteczek zmienia swą prędkość, a więc przechodzi do innego przedziału. Na ich miejsce w tym samym czasie wchodzi taka sama liczba cząsteczek innych, takich, które poprzednio miały prędkości leżące poza granicami v i $v+dv$. Jest to słuszne jednak jedynie wtedy, gdy ogólna liczba N cząsteczek badanych jest bardzo duża.



Rys. 16.3

Wyprowadzenia prawa rozkładu prędkości nie będziemy podawali. Wymieniliśmy jedynie podstawowe założenia, na jakich opiera się dowód Maxwella. Przeanalizujemy jeszcze krzywą spełniającą prawo rozkładu prędkości (rys. 16.3) i na zakończenie podamy matematyczną postać tego prawa. Na osi odciętych odłożone są prędkości cząsteczek poczynając od wartości zerowej. Rzędne przedstawiają stosunek dN/Ndv, tzn. ułamek ogólnej liczby cząsteczek przypadający na jednostkowy przedział prędkości.

Można by też wykreślić krzywą Maxwella odkładając na osi rzędnych stosunek dN/dv. Otrzymalibyśmy krzywą o tym samym charakterze, tylko o skali pionowej N razy powiększonej. Zastanówmy się, co przedstawiałoby przy takich współrzędnych pole prostokąta o podstawie dv i wysokości dN/dv. Iloczyn $(dN/dv)\,dv$ daje liczbę dN. Innymi słowy, pole prostokąta wyrażałoby liczbę cząsteczek o prędkościach zawartych w granicach v i $v+dv$. Stopniowo przesuwając się od prędkości najmniejszych do największych moglibyśmy podzielić całe pole zawarte między krzywą a osią odciętych na nieskończenie wiele takich prostokątów, których pola wyrażałyby liczby cząsteczek dN_1, dN_2, dN_3 itd. o prędkościach zawartych w odpowiednich granicach. Suma dN_1+ $+dN_2+dN_3+\ldots+dN_n$, gdy $n \to \infty$, równa się N, a więc pole całkowite zawarte między krzywą a osią odciętych przedstawiałoby ogólną liczbę N cząsteczek gazu. Wiemy już, że ogólna liczba cząsteczek gazu zawartych w warunkach normalnych w objętości 1 cm^3 jest rzędu 10^{19}. Stąd wniosek, że rozmiary naszej krzywej wymagałyby bardzo dużej skali i byłyby zależne od rozmiarów naczynia i warunków, w jakich gaz się znajduje. Dlatego wygodnie jest na osi rzędnych odkładać nie dN/dv, lecz $dN/N\,dv$. Przy takich współrzędnych pole zakreskowane na rys. 16.3 odpowiadające nieskończenie małej zmianie prędkości dv wyraża ułamek dN/N ogólnej liczby cząsteczek, przypadający na przedział prędkości v i $v+dv$. A zatem całe pole zawarte między osią odciętych i krzywą będzie sumą

$$\frac{dN_1}{N}+\frac{dN_2}{N}+\frac{dN_3}{N}+\ldots+\frac{dN_n}{N}=\frac{dN_1+dN_2+dN_3+\ldots+dN_n}{N}=1,$$

przy czym $n \to \infty$. W ten sposób zmniejsza się skala naszego wykresu i rozmiary krzywej stają się niezależne od liczby N cząsteczek zawartych w badanej masie gazu. W dalszym ciągu jednak obowiązuje założenie, że liczba N jest bardzo duża.

Z wykresu krzywej Maxwella wyciągamy następujące wnioski. Cząsteczki gazu mają różne wartości prędkości, od bardzo małych, bliskich zera, do bardzo dużych. Z wykresu widać, że zarówno bardzo małym, jak i bardzo dużym prędkościom odpowiadają małe rzędne, tzn. zarówno pierwsze, jak i drugie występują niezmiernie rzadko.

Z istnienia maksimum krzywej wnioskujemy, że pewnej określonej wartości prędkości odpowiada największa liczba cząsteczek. Jest to tzw. prędkość najprawdopodobniejsza $\hat{v}$. Duży procent cząsteczek ma prędkości mało różniące się od najprawdopodobniejszej. W tabeli 16.1 mamy procentowy rozkład prędkości cząsteczek tlenu w temperaturze 273,15K (0°C).

Sumując wartości liczbowe prędkości wszystkich cząsteczek i dzieląc wynik przez liczbę cząsteczek N znajdujemy tzw. *prędkość średnią* $\bar{v}$:

$$\bar{v}=\frac{v_1+v_2+v_3+\ldots+v_N}{N}.$$

Z przebiegu krzywej, a dokładniej mówiąc z faktu, że krzywa nie jest symetryczna względem osi pionowej przeprowadzonej przez wartość prędkości najprawdopodobniejszej $\hat{v}$, wnioskujemy, że prędkość średnia $\bar{v} \neq \hat{v}$.

Wspomniana oś pionowa dzieli pole zawarte między krzywą a osią odciętych na dwie części, przy czym większa jest część pola leżąca po stronie prędkości większych.

Tabela 16.1

Procentowy rozkład prędkości cząsteczek tlenu w temperaturze 273,15K (0°C)

Granice prędkości m/s	Liczba cząsteczek %
$v < 100$	1,4
100–200	8,1
200–300	16,7
300–400	21,5
400–500	20,3
500–600	15,1
600–700	9,2
$v > 700$	7,7

Istnieje zatem przewaga w liczbie cząsteczek o prędkościach większych od najprawdopodobniejszej, czyli $\bar{v} > \hat{v}$.

Dla ilustracji zależności matematycznej między wymienionymi wielkościami podajemy wzór Maxwella, odpowiadający wykreślonej krzywej rozkładu prędkości:

$$\frac{dN}{N} = \frac{4}{\sqrt{\pi}} \frac{v^2}{\hat{v}^3} e^{-\frac{v^2}{\hat{v}^2}} dv,$$

gdzie N oznacza ogólną liczbę cząsteczek, dN — liczbę cząsteczek odpowiadającą przedziałowi prędkości v i $v+dv$, $\hat{v}$ — prędkość najprawdopodobniejszą, e — podstawę logarytmów naturalnych = 2,718...

16.4. Ciśnienie gazu z punktu widzenia teorii kinetycznej

Jeden z podstawowych wzorów teorii kinetycznej wyraża związek między ciśnieniem gazu a wielkościami cząsteczkowymi. Wyprowadzimy taki wzór stosując pewne założenia upraszczające:

1. *Gaz jest zawarty w naczyniu kulistym o promieniu r.*

2. *Gaz jest tak rozrzedzony, że można brać pod uwagę tylko zderzenia ze ściankami naczynia zaniedbując zderzenia międzycząsteczkowe.*

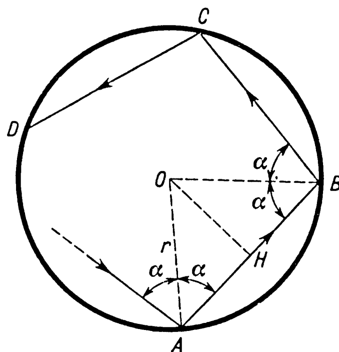
Prześledźmy bieg jednej cząsteczki o masie m i prędkości $\mathbf{v}$, poruszającej się w płaszczyźnie rys. 16.4. Cząsteczka ta uderza w punkcie A o ściankę naczynia. Kierunek prędkości $\mathbf{v}$ tworzy z normalną do ścianki naczynia, wystawioną w punkcie A (czyli z promieniem OA), kąt α.

Zakładamy, że cząsteczki podczas zderzeń zachowują się jak kulki doskonale sprężyste. Na podstawie prawa, że kąt padania równa się kątowi odbicia, kąt OAB jest również α. Te same kąty padania i odbicia występują przy zderzeniach ze ścianką w punktach B, C, D i dalszych. Z praw odbicia wynika też, że kolejne odcinki toru AB, BC itd. leżą w jednej płaszczyźnie. Zmiana płaszczyzny ruchu mogłaby wystąpić

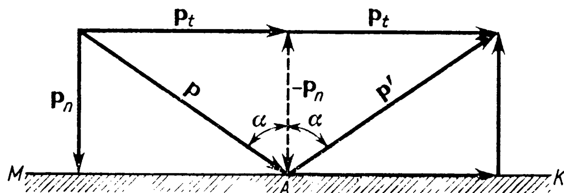
po zderzeniu cząsteczki badanej z inną, ale tego, zgodnie z założeniem 2, nie bierzemy pod uwagę. Ze stosunków geometrycznych wynika, że $AB = BC = CD$ itd.

Obliczenie średniego ciśnienia p, działającego na powierzchnię S, wymaga znajomości ogólnej siły F, działającej na tę powierzchnię. Z drugiej zasady dynamiki wiadomo, że

$$F = \frac{d(mv)}{dt},$$



Rys. 16.4



Rys. 16.5

a zatem obliczenie siły można sprowadzić do wyznaczenia zmiany pędu zachodzącej w jednostce czasu. Zaczniemy od obliczenia *zmiany pędu pojedynczej cząsteczki przy jednym jej zderzeniu ze ścianką naczynia.* Rozważmy pomocniczo zmianę pędu kulki doskonale sprężystej o masie m, uderzającej o nieruchomą ściankę MK (rys. 16.5) (stanowiącą część płaszczyzny stycznej do ściany naczynia w punkcie A) z prędkością $\mathbf{v}$ pod kątem α. Kulka ta ma w momencie zderzenia pęd $\mathbf{p}(\mathbf{p} = m\mathbf{v})$, skierowany zgodnie z prędkością. Rozkładamy pęd $\mathbf{p}$ na składową styczną $\mathbf{p}_t$, równą liczbowo $mv \sin\alpha$, i składową normalną $\mathbf{p}_n$, równą liczbowo $mv \cos\alpha$. Podczas zderzenia składowa styczna nie ulega zmianie, natomiast składowa normalna zmienia swój znak na przeciwny (odpowiada to zmianie kierunku składowej normalnej prędkości w chwili odbicia). A zatem zmiana pędu, powstająca podczas odbicia, równa się różnicy wartości liczbowych składowych normalnych pędu przed i po odbiciu:

$$mv \cos\alpha - (-mv \cos\alpha) = 2mv \cos\alpha.$$

Przenosząc ten wynik na rozważane zagadnienie ruchu cząsteczek gazowych stwierdzimy, że jedna cząsteczka w czasie pojedynczego uderzenia o ściankę doznaje zmiany pędu $2mv\cos\alpha$. Chcąc określić *całkowitą siłę* działającą na ściankę naczynia *pochodzącą od jednej cząsteczki* musimy obliczyć zmianę jej·pędu w czasie 1 s. Do tego potrzebna jest znajomość liczby n uderzeń cząsteczki o ściankę w czasie 1 s:

$$n = \frac{v}{AB}.$$

Ale

$$AB = 2AH = 2r\cos\alpha.$$

Zmiana pędu odpowiadająca n uderzeniom (w czasie 1 s), będąca miarą siły oddziaływania jednej cząsteczki na ściankę naczynia, wyraża się wzorem

$$f = 2mv\cos\alpha\cdot n = \frac{2mv\cos\alpha\cdot v}{2r\cos\alpha} = \frac{mv^2}{r}.$$

W naczyniu mamy nie jedną, ale N cząsteczek o jednakowych masach, jeśli gaz jest jednorodny, o różnych prędkościach $v_1, v_2, v_3, \ldots, v_N$. Całkowita siła F, działająca na powierzchnię wewnętrzną naczynia, pochodząca od wszystkich cząsteczek, jest sumą oddziaływań poszczególnych cząsteczek:

$$F = \frac{\sum_{i=1}^{N} mv_i^2}{r}.$$

Pamiętając, że ciśnienie $p = F/S$, a $S = 4\pi r^2$, znajdujemy

$$p = \frac{\sum_{i=1}^{N} mv_i^2}{4\pi r^3}, \qquad \text{skąd} \qquad p = \frac{\sum_{i=1}^{N} mv_i^2}{3\cdot\frac{4}{3}\pi r^3},$$

czyli

$$p = \frac{\sum_{i=1}^{N} mv_i^2}{3V}. \tag{16.1}$$

Aby lepiej zdać sobie sprawę z treści tego równania, przepiszemy licznik w innej postaci:

$$\sum_{i=1}^{N} mv_i^2 = m(v_1^2+v_2^2+\ldots+v_N^2).$$

Mnożąc i dzieląc to wyrażenie przez $2N$ otrzymamy

$$\sum_{i=1}^{N} mv_i^2 = \frac{2Nm}{2}\,\frac{v_1^2+v_2^2+\ldots+v_N^2}{N}.$$

Stosunek sumy kwadratów prędkości poszczególnych cząsteczek do ogólnej ich liczby

nazywamy *średnim kwadratem prędkości* i oznaczamy symbolem $\overline{v^2}$ (należy tę wielkość odróżniać od kwadratu prędkości średniej $\bar{v}^2$), czyli

$$\overline{v^2} = \frac{v_1^2+v_2^2+\ldots+v_N^2}{N}.$$

Po uwzględnieniu tej zależności w równaniu poprzednim otrzymujemy

$$\sum_{i=1}^{N} mv_i^2 = 2N\frac{m\overline{v^2}}{2}.$$

Wyrażenie $m\overline{v^2}/2 = \bar{E}_k$ jest średnią energią kinetyczną cząsteczki w jej ruchu postępowym. Iloczyn tej wielkości i N daje energię kinetyczną wszystkich cząsteczek gazu:

$$E_k = N\frac{m\overline{v^2}}{2}.$$

A zatem wzór na ciśnienie przyjmuje postać

$$p = \tfrac{2}{3}\frac{E_k}{V}. \tag{16.2}$$

Celem uproszczenia rozważań matematycznych wprowadziliśmy pewne założenia. Można jednak udowodnić, że do gazów o większym stopniu zgęszczenia, zawartych w naczyniach dowolnych kształtów, stosuje się też wzór ostateczny tej samej postaci.

16.5. Obliczanie prędkości średniej kwadratowej

Jedną z form przejściowych wzoru na ciśnienie jest wyrażenie

$$p = \tfrac{1}{3}\frac{Nm\overline{v^2}}{V}, \tag{16.3}$$

ale $Nm = M$ równa się całkowitej masie gazu. Stosunek M/V określa gęstość gazu ϱ. A zatem

$$p = \tfrac{1}{3}\varrho\overline{v^2}, \tag{16.4}$$

skąd

$$\overline{v^2} = \frac{3p}{\varrho},$$

a prędkość średnia kwadratowa, czyli pierwiastek kwadratowy ze średniego kwadratu wynosi

$$\sqrt{\overline{v^2}} = \sqrt{\frac{3p}{\varrho}}.$$

Poznaliśmy zatem trzy różne prędkości charakteryzujące ruch cząsteczek gazowych, a mianowicie prędkość średnią $\bar{v}$, prędkość średnią kwadratową $\sqrt{\overline{v^2}}$ i prędkość najprawdopodobniejszą $\hat{v}$. Okazuje się, że ich wartości liczbowe spełniają następującą nierówność:

$$\hat{v} < \bar{v} < \sqrt{\overline{v^2}}.$$

Można udowodnić, że prędkość średnia

$$\bar{v} = \frac{2\hat{v}}{\sqrt{\pi}} \approx 1{,}13\hat{v},$$

a prędkość średnia kwadratowa

$$\sqrt{\overline{v^2}} = \hat{v}\sqrt{\tfrac{3}{2}} \approx 1{,}21\hat{v}.$$

16.6. Pojęcie temperatury z punktu widzenia teorii kinetycznej

Przez przekształcenie wzoru na ciśnienie można wyrazić iloczyn pV jako

$$pV = \tfrac{2}{3}E_k,$$

czyli

$$pV = \tfrac{2}{3}N\frac{m\overline{v^2}}{2}.$$

Załóżmy, że badana ilość gazu odpowiada jednemu molowi. Wtedy N odpowiada liczbie Avogadra N_A, a V wyraża V_m. Z teorii gazów doskonałych wiadomo, że w odniesieniu do jednego mola obowiązuje równanie Clapeyrona

$$pV_m = RT,$$

a zatem

$$\tfrac{2}{3}N_A\frac{m\overline{v^2}}{2} = RT.$$

Stąd średnia energia kinetyczna pojedynczej cząsteczki

$$\frac{m\overline{v^2}}{2} = \frac{3}{2}\frac{R}{N_A}T.$$

Stosunek R/N_A (uniwersalnej stałej gazowej do stałej Avogadra) nosi nazwę *stałej Boltzmanna* i oznaczany jest literą k:

$$k = \frac{R}{N_A} = \frac{8{,}314}{6{,}02\cdot 10^{23}} = 1{,}380\cdot 10^{-23}\ \mathrm{J/K}.$$

Z równania

$$\frac{m\overline{v^2}}{2} = \tfrac{3}{2}kT \qquad (16.5)$$

wnioskujemy, że *średnia energia kinetyczna ruchu postępowego cząsteczek jest tylko funkcją temperatury bezwzględnej. Nie zależy od rodzaju gazu ani od jego ciśnienia.*

Równanie (16.5) przedstawia kinetyczne określenie temperatury. Pamiętać jednak trzeba, że *temperatura jest wielkością statystyczną*, tzn. kinetyczne określenie temperatury ma sens tylko w odniesieniu do wielkiej liczby cząsteczek. Gdyby w pewnej przestrzeni znajdowało się zaledwie kilka czy kilkanaście cząsteczek, to wartość ich średniej energii kinetycznej zmieniałaby się od chwili do chwili. W takim układzie nie można byłoby określać temperatury za pośrednictwem średniej energii kinetycznej.

Załóżmy, że dwa różne gazy mają jednakowe temperatury T. Cząsteczki pierwszego gazu mają masy m_1, cząsteczki drugiego gazu — masy m_2. Wobec równości temperatur średnie energie kinetyczne obu gazów są jednakowe:

$$\frac{m_1 \overline{v_1^2}}{2} = \frac{m_2 \overline{v_2^2}}{2},$$

skąd

$$\frac{m_1}{m_2} = \frac{\overline{v_2^2}}{\overline{v_1^2}}.$$

W tej samej temperaturze średnie kwadraty prędkości dwóch różnych gazów są odwrotnie proporcjonalne do mas ich cząsteczek. Cząsteczki lżejsze mają prędkości średnie kwadratowe większe i odwrotnie.

16.7. Prawa charakteryzujące przemiany gazowe w ujęciu teorii kinetycznej

16.7.1. *Prawo Boyle'a–Mariotte'a z punktu widzenia teorii kinetycznej gazów.* Powróćmy do równania

$$pV = \tfrac{2}{3} N \frac{m\overline{v^2}}{2}.$$

Równanie Boyle'a–Mariotte'a odnosi się do przemian izotermicznych określonej masy gazu. W stałej temperaturze średnia energia kinetyczna cząsteczek jest niezmienna. Jeśli masa jest określona, to i liczba cząsteczek N zawartych w tej masie jest stała. Czynniki iloczynu po prawej stronie równania są zatem stałe, a więc i iloczyn $pV = \text{const}$, co odpowiada znanemu prawu Boyle'a–Mariotte'a.

16.7.2. *Prawo Avogadra z punktu widzenia teorii kinetycznej gazów. Prawo Avogadra mówi, że w jednakowych objętościach różnych gazów mierzonych pod tym samym ciśnieniem i w tej samej temperaturze znajduje się jednakowa liczba cząsteczek.* Prawo to można bezpośrednio odczytać ze wzoru na ciśnienie gazu. Załóżmy, że w objętości V w temperaturze T znajduje się N_1 cząsteczek gazu wywierającego ciśnienie p:

$$p = \tfrac{2}{3} \frac{N_1}{V} \frac{m_1 \overline{v_1^2}}{2}.$$

W odniesieniu do innego gazu o tej samej temperaturze, zamkniętego w tej samej objętości i wywierającego to samo ciśnienie, obowiązuje równanie

$$p = \tfrac{2}{3} \frac{N_2}{V} \frac{m_2 \overline{v_2^2}}{2}.$$

Wobec równości p, V oraz średnich energii kinetycznych cząsteczek obu gazów (jednakowe temperatury) stwierdzamy, że

$$N_1 = N_2,$$

co jest treścią prawa Avogadra.

16.7.3. *Prawo Daltona z punktu widzenia teorii kinetycznej gazów.* Przy wyprowadzaniu wzoru na ciśnienie gazu odrzućmy warunek, że gaz jest jednorodny. Niech

naczynie będzie wypełnione mieszaniną kilku gazów. Masy cząsteczek oraz liczbę cząsteczek składników mieszaniny oznaczmy odpowiednio przez m_1 i N_1, m_2 i N_2 itd. Wtedy zgodnie z wzorem (16.1)

$$p = \frac{1}{3V}\left(\sum_{i=1}^{N_1} m_1 v_{1i}^2 + \sum_{j=1}^{N_2} m_2 v_{2j}^2 + \sum_{k=1}^{N_3} m_3 v_{3k}^2 + \dots\right),$$

gdyż cząsteczki wszystkich gazów biorą udział w wywieraniu ciśnienia na ścianki naczynia. Wielkości $v_{11}, v_{12}, v_{13}, \dots, v_{1N_1}$ oznaczają prędkości poszczególnych cząsteczek pierwszego składnika mieszaniny. Podobnie $v_{21}, v_{22}, v_{23}, \dots, v_{2N_2}$ oznaczają prędkości cząsteczek składnika drugiego itd. Ale

$$\sum_{i=1}^{N_1} m_1 v_{1i}^2 = \frac{N_1(m_1 v_{11}^2 + m_1 v_{12}^2 + \dots + m_1 v_{1N_1}^2)}{N_1} = N_1 m_1 \overline{v_1^2}.$$

Analogicznie

$$\sum_{j=1}^{N_2} m_2 v_{2j}^2 = \frac{N_2(m_2 v_{21}^2 + m_2 v_{22}^2 + \dots + m_2 v_{2N_2}^2)}{N_2} = N_2 m_2 \overline{v_2^2}$$

itd., a zatem

$$p = \frac{1}{3V}(N_1 m_1 \overline{v_1^2} + N_2 m_2 \overline{v_2^2} + N_3 m_3 \overline{v_3^2} + \dots),$$

co można przepisać w postaci

$$p = \frac{1}{3}\frac{N_1 m_1 \overline{v_1^2}}{V} + \frac{1}{3}\frac{N_2 m_2 \overline{v_2^2}}{V} + \frac{1}{3}\frac{N_3 m_3 \overline{v_3^2}}{V} + \dots$$

Składniki tej sumy określają ciśnienia cząstkowe $p_1, p_2, p_3, \dots$, jakie wywierałyby poszczególne gazy wchodzące w skład mieszaniny zajmując każdy z osobna całą objętość V:

$$p = p_1 + p_2 + p_3 + \dots,$$

co odpowiada prawu Daltona.

16.8. Zasada ekwipartycji energii

Przede wszystkim ustalimy znaczenie wyrazu ekwipartycja. Pochodzi on od wyrazów łacińskich *aequus* — równy, *pars* — część. Zasada ekwipartycji energii podana przez Maxwella mówi o równym podziale energii. Zanim przejdziemy do dokładnego ujęcia treści tej zasady zapoznamy się z pojęciem stopni swobody. *Przez liczbę stopni swobody jakiegoś ciała, cząsteczki czy atomu rozumiemy liczbę zmiennych niezależnych, charakteryzujących położenie badanego układu w przestrzeni.* Można też liczbę stopni swobody określać przez podanie liczby niezależnych ruchów, jakie może wykonywać dany układ. Tak np. położenie punktu materialnego lub cząsteczki jednoatomowej w przestrzeni charakteryzują trzy niezależne współrzędne x, y, z w układzie Kartezjusza. Cząsteczki jednoatomowe, np. cząsteczki helu, argonu itp., mają zatem trzy stopnie swobody.

Cząsteczkę złożoną z dwóch atomów można w pierwszym przybliżeniu traktować jako układ dwóch punktów materialnych pozostających w stałej odległości. Położenie w przestrzeni każdego z tych punktów jest określone trzema współrzędnymi, ale wobec warunku stałej ich odległości mamy jedno równanie wiążące ze sobą współrzędne. Niezależnych współrzędnych pozostaje pięć. Cząsteczki dwuatomowe (H_2, O_2 itp.) o stałej odległości między atomami mają właśnie pięć stopni swobody.

Trzy punkty pozostające w stałych odległościach od siebie mogą stanowić odpowiednik sztywnej cząsteczki trójatomowej. Każdy z punktów jest określony trzema współrzędnymi. Razem więc mamy dziewięć współrzędnych. Warunek stałych odległości między tymi punktami (nie leżącymi na jednej prostej) narzuca nam trzy równania wiążące poszczególne współrzędne. Z ogólnej liczby dziewięciu współrzędnych pozostaje tylko sześć współrzędnych niezależnych, a więc układ taki charakteryzuje 6 stopni swobody.

Analogiczne rozumowanie dla sztywno związanych cząsteczek 4-, 5- i więcej atomowych pozwoli stwierdzić, że i takie cząsteczki będą miały po 6 stopni swobody. Trzeba tylko dobrze określić liczbę równań określających stałe odległości między wszystkimi możliwymi parami atomów w rozpatrywanej cząsteczce.

Rozpatrując zagadnienie stopni swobody ze względu na rodzaje możliwych ruchów można łatwo stwierdzić, że cząsteczka jednoatomowa może wykonywać trzy niezależne ruchy postępowe wzdłuż trzech wzajemnie prostopadłych osi, ma zatem trzy stopnie swobody. Cząsteczka dwuatomowa prócz trzech ruchów postępowych może wykonywać jeszcze dwa ruchy obrotowe dokoła dwóch osi wzajemnie prostopadłych i prostopadłych do linii łączącej oba atomy traktowane jako punkty materialne. Cząsteczka taka ma więc 5 stopni swobody. Ruchu obrotowego dokoła osi cząsteczki, tzn. dokoła linii łączącej oba atomy, oczywiście nie uwzględniamy.

Cząsteczka trójatomowa, której atomy nie leżą na jednej prostej, wykonywać może trzy ruchy postępowe i trzy ruchy obrotowe dokoła trzech wzajemnie prostopadłych osi, ma zatem 6 stopni swobody. Podobnie cząsteczka wieloatomowa ze względu na 6 możliwych ruchów ma 6 stopni swobody.

Zasada ekwipartycji energii mówi o równomiernym podziale średniej energii kinetycznej na poszczególne stopnie swobody. Według tej zasady na każdy stopień swobody cząsteczki (niezależnie od rodzaju ruchu i niezależnie od właściwości chemicznych badanej substancji) przypada średnia energia kinetyczna w ilości $\frac{1}{2}kT$, gdzie k oznacza stałą Boltzmanna, T — temperaturę bezwzględną.

Rozważmy to dla cząsteczki jednoatomowej, poruszającej się w przestrzeni z prędkością v. Cząsteczka taka, jak już wiemy, ma trzy stopnie swobody. Niech początek układu odniesienia (x, y, z) pokrywa się z początkiem wektora $\mathbf{v}$. Rozłóżmy prędkość na trzy składowe równoległe do kierunków osi współrzędnych, a mianowicie na $\mathbf{v}_x$, $\mathbf{v}_y$ i $\mathbf{v}_z$ (rys. 16.6). Ujmując zależności wektorowo napiszemy

$$\mathbf{v} = \mathbf{v}_x+\mathbf{v}_y+\mathbf{v}_z, \qquad \overrightarrow{OB} = \mathbf{v}_x+\mathbf{v}_y.$$

Analitycznie zaś, uwzględniając, że

$$OB^2 = v_x^2+v_y^2,$$

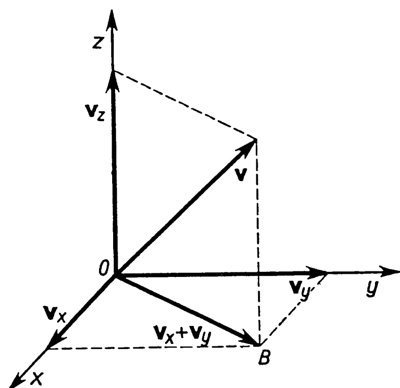
otrzymujemy

$$v^2 = v_x^2+v_y^2+v_z^2.$$

Pomnóżmy całe to równanie przez $m/2$, tzn. przez połowę masy pojedynczej cząsteczki

$$\frac{mv^2}{2} = \frac{mv_x^2}{2}+\frac{mv_y^2}{2}+\frac{mv_z^2}{2}.$$

Poszczególne składniki tego równania wyrażają następujące energie kinetyczne:
$mv^2/2$ — całkowitą energię kinetyczną cząsteczki,
$mv_x^2/2$, $mv_y^2/2$ i $mv_z^2/2$ — energie kinetyczne związane odpowiednio z ruchami składowymi w kierunkach osi Ox, Oy i Oz.



Rys. 16.6

Nie mamy żadnych podstaw do przypuszczenia, że dla dowolnie wybranej pojedynczej cząsteczki energie $\frac{mv_x^2}{2}$, $\frac{mv_y^2}{2}$ i $\frac{mv_z^2}{2}$, odpowiadające ruchom składowym w kierunku osi Ox, Oy i Oz, są jednakowe. Stosując jednak nasze rozumowanie do bardzo dużej liczby N cząsteczek otrzymamy:

dla pierwszej cząsteczki $mv_1^2/2 = mv_{1x}^2/2+mv_{1y}^2/2+mv_{1z}^2/2$,

dla drugiej cząsteczki $mv_2^2/2 = mv_{2x}^2/2+mv_{2y}^2/2+mv_{2z}^2/2$,

. . . .

dla N-tej cząsteczki $mv_N^2/2 = mv_{Nx}^2/2+mv_{Ny}^2/2+mv_{Nz}^2/2$.

Sumujemy stronami wszystkie wyrażenia i dzielimy przez N. Po lewej stronie otrzymujemy średnią energię kinetyczną, przypadającą na pojedynczą cząsteczkę w ruchu postępowym (energia translacji), po prawej zaś mamy sumę trzech średnich energii kinetycznych, odpowiadających ruchom składowym w kierunku osi Ox, Oy i Oz:

$$\overline{E} = \overline{E}_x+\overline{E}_y+\overline{E}_z.$$

Wiemy z rozważań poprzednich, że przy dostatecznie dużej liczbie cząsteczek biorących udział w ruchu cieplnym można ten ruch uważać za doskonale chaotyczny: ża-

den kierunek ruchu nie jest uprzywilejowany. Nie może występować przewaga którejkolwiek z trzech średnich energii kinetycznych, czyli

$$\bar{E}_x = \bar{E}_y = \bar{E}_z = \tfrac{1}{3}\bar{E}.$$

W rozważaniach poprzednich doszliśmy też do wniosku, że średnia energia kinetyczna ruchu postępowego cząsteczki jednoatomowej

$$\bar{E}_k = \tfrac{3}{2}kT.$$

Stąd wniosek, że

$$\bar{E}_x = \bar{E}_y = \bar{E}_z = \tfrac{1}{2}kT,$$

czyli na każdy stopień swobody przypada na poszczególną cząsteczkę średnia energia kinetyczna równa połowie iloczynu kT. Udowodniliśmy to dla cząsteczki o trzech stopniach swobody, tzn. cząsteczki jednoatomowej, odbywającej tylko ruch postępowy. Zasada Maxwella odnosi się jednak do wszystkich rodzajów ruchów, przydzielając każdej cząsteczce na każdy stopień swobody średnią energię kinetyczną w ilości $\frac{1}{2}kT$.

Znając zasadę ekwipartycji energii i liczbę stopni swobody f, charakteryzujących cząsteczki badanego gazu, można obliczyć energię kinetyczną np. 1 mola ($N_A = 6{,}02 \cdot 10^{23}$ cząsteczek), zwaną też energią molową E_m:

$$E_m = f \cdot \tfrac{1}{2}kTN_A = \frac{f}{2}\frac{R}{N_A}TN_A = \frac{fTR}{2}. \tag{16.6}$$

W przypadku *gazu doskonałego* (tzn. gdy *zaniedbujemy oddziaływania międzycząsteczkowe*) *energia kinetyczna* E_k *jest równocześnie całkowitą energią wewnętrzną* U:

$$U = E_k.$$

Całkowita energia wewnętrzna U *gazu rzeczywistego* nie równa się jego energii kinetycznej, gdyż międzycząsteczkowe, nieznaczne zresztą, oddziaływania decydują o istnieniu energii potencjalnej E_p. Wartość energii wewnętrznej gazu rzeczywistego

$$U = E_k + E_p$$

zależy w dużym stopniu od zgęszczenia gazu rzeczywistego. Pod małymi ciśnieniami, gdy cząsteczki gazu znajdują się w dużych odległościach od siebie, energię potencjalną można zaniedbywać przyjmując, że energia wewnętrzna $U = E_k$ (przybliżenie prowadzące do modelu gazu doskonałego).

W odniesieniu do 1 mola gazu doskonałego można też mówić o energii wewnętrznej molowej U_m. Zgodnie z (16.6) skoro

$$U_m = E_m,$$

to

$$U_m = \frac{fRT}{2}. \tag{16.7}$$

Na zakończenie jeszcze jedna uwaga. Ustaliliśmy liczbę stopni swobody f:

dla cząsteczek jednoatomowych $f = 3$,

dla cząsteczek dwuatomowych $f = 5$,

dla cząsteczek trój- i wieloatomowych $f = 6$.

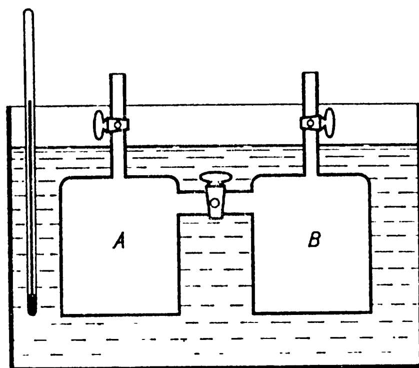
Zagadnienie liczby stopni swobody komplikuje się w temperaturach wyższych. Gdy energie poszczególnych czasteczek są duże, zderzenia mogą wywoływać powstawanie nowego rodzaju ruchu. Zderzenia pobudzają atomy w cząsteczkach do drgań, czyli oscylacji, wzrasta więc liczba niezależnych ruchów, wzrasta liczba stopni swobody. Zagadnienie staje się tym bardziej skomplikowane, że ruch oscylacyjny atomów występuje nie od razu we wszystkich cząsteczkach w pewnej określonej temperaturze, ale w miarę wzrostu temperatury pojawia się w coraz większej liczbie cząsteczek. Trzeba też pamiętać, że z oscylacjami wiąże się nie tylko energia kinetyczna, lecz także energia potencjalna, zależna od kwadratu wychylenia od położenia równowagi.

Innymi słowy, istnieją pewne trudności podczas stosowania zasady ekwipartycji energii. Zmianom temperatury towarzyszy zmiana liczby stopni swobody danego ciała. Ze wzrostem temperatury liczba stopni swobody rośnie. Gdy temperatura zbliża się do zera skali Kelvina, zamiera ruch obrotowy cząsteczek, tzn. cząsteczki wieloatomowe pod względem liczby stopni swobody schodzą do rzędu cząsteczek jednoatomowych.

Ostatnio zasadę ekwipartycji energii Maxwella uważa się za zasadę odnoszącą się do takich stanów ustalonych, w których wszystkie cząsteczki mają tę samą liczbę stopni swobody. Nie stosuje się ona natomiast do stanów przejściowych, w których zmienia się stopniowo liczba cząsteczek o określonej liczbie stopni swobody.

16.9. Energia wewnętrzna gazu doskonałego jako funkcja temperatury

Ustaliliśmy ostatnio, że energia wewnętrzna gazu doskonałego $U = E_k$ ruchu molekularnego. Nasuwa się pytanie, czy energia wewnętrzna gazu doskonałego jest funkcją objętości, ciśnienia i temperatury. Odpowiedź na to pytanie daje doświadczenie Gay-Lussaca (rys. 16.7). Dwa zbiorniki A i B połączone są przewodem zaopatrzonym w za-



Rys. 16.7

wór przelotowy. Całość zanurzona jest w naczyniu kalorymetrycznym wypełnionym wodą. W zbiorniku A mamy próżnię, w zbiorniku B zamknięta jest pewna ilość gazu w objętości V pod ciśnieniem p. Woda w kalorymetrze i oba zbiorniki mają temperaturę T. Otwieramy zawór przelotowy. Gaz bez pokonywania jakiegokolwiek oporu, a więc bez wykonywania pracy, zajmuje łączną objętość zbiorników A i B.

Ze wzrostem objętości maleje ciśnienie gazu. Temperatura jednak, jak wynika z pomiarów, pozostaje niezmieniona.

Analizujemy raz jeszcze doświadczenie: praca nie została włożona w układ, ani oddana na zewnątrz. Nie było straty ani zysku ciepła. Stąd wniosek, że energia wewnętrzna nie uległa zmianie.

Podczas doświadczenia zmieniły się objętość i ciśnienie, niezmieniona została temperatura. Stąd dalszy wniosek, że energia wewnętrzna nie zależy ani od ciśnienia, ani od objętości, czyli nie jest funkcją odległości międzycząsteczkowych, a jest tylko funkcją temperatury.

Dokładniejsze pomiary przeprowadzone później przez Joule'a i Thomsona wykazały, że gazy rzeczywiste rozprężając się zmieniają w małym stopniu swą temperaturę. Jest to tzw. efekt Joule'a–Thomsona. Podczas doświadczeń przeprowadzanych w temperaturach bliskich 0°C większość gazów, między innymi tlen i azot, oziębiała się, wodór zaś, hel i neon lekko się ogrzewały. Jeśli jednak rozprężanie odbywało się w wysokich temperaturach, to zmiany temperatury stawały się mniejsze. Wiemy, że w wysokich temperaturach właściwości gazów rzeczywistych mało różnią się od właściwości gazów doskonałych. A zatem wniosek poprzednio wyciągnięty, że energia wewnętrzna jest tylko funkcją temperatury przyjmujemy za słuszny i dokładnie obowiązujący tylko dla gazu doskonałego.

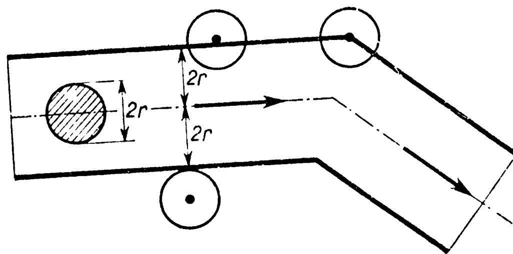
Zależność energii wewnętrznej gazu doskonałego od temperatury dla jednego mola ustaliliśmy już poprzednio (por. § 16.7). Przyrost energii wewnętrznej molowej ΔU_m gazu doskonałego wyraża się wzorem:

$$\Delta U_m = \tfrac{1}{2} f R \Delta T. \qquad (16.8)$$

16.10. Średnia droga swobodna cząsteczek

Średnią drogę swobodną cząsteczki określiliśmy jako średnią długość odcinka prostoliniowego przebiegu między dwoma zderzeniami.

Załóżmy, że w przestrzeni wypełnionej gazem wśród bardzo licznych cząsteczek nieruchomych porusza się jedna z prędkością $\bar{v}$. Wszystkie cząsteczki potraktujemy jako



Rys. 16.8

kulki o promieniach r. Chwila zastanowienia wystarczy do stwierdzenia za pomocą rys. 16.8, że biegnąca cząsteczka w ciągu jednej sekundy zderzy się ze wszystkimi cząsteczkami, których środki znajdują się w walcu o promieniu podstawy równym $2r$,

a wysokości równej $\bar{v}$. Zakładając, że w jednostce objętości znajduje się n cząsteczek gazu, otrzymujemy liczbę cząsteczek zawartych we wspomnianym walcu mnożąc n przez objętość walca $\pi(2r)^2\bar{v}$. Tyleż wynosi liczba zderzeń ν' w jednostce czasu:

$$\nu' = 4\pi r^2 n\bar{v}.$$

Z rozważań matematycznych wynika, że uwzględniając ruch pozostałych cząsteczek musimy wprowadzić do tego iloczynu poprawkę w postaci czynnika $\sqrt{2}$. W tych warunkach liczba zderzeń ν w jednostce czasu wyrazi się wzorem:

$$\nu = 4\sqrt{2}\pi r^2 n\bar{v}.$$

Całkowita długość drogi przebytej przez cząsteczkę w jednostce czasu jest liczbowo równa $\bar{v}$. Średnia długość odcinka między dwoma kolejnymi zderzeniami, czyli średnia droga swobodna l, wynosi zatem

$$l = \frac{\bar{v}}{\nu} = \frac{\bar{v}}{4\sqrt{2}\pi r^2 n\bar{v}} = \frac{1}{4\sqrt{2}\pi r^2 n}.$$

W tabeli 16.2 podane jest zestawienie wartości liczbowych kilku charakterystycznych wielkości z teorii kinetycznej dla paru gazów rzeczywistych w warunkach normalnych.

Tabela 16.2

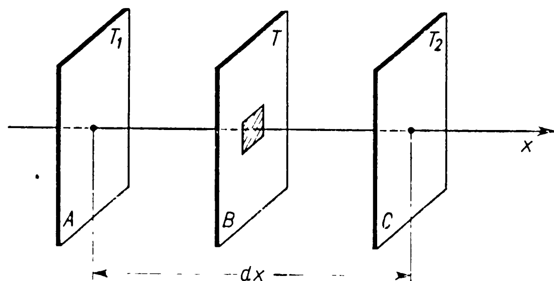
Zestawienie charakterystycznych wielkości z teorii kinetycznej dla kilku gazów w warunkach normalnych

Rodzaj gazu	Prędkość średnia $\bar{v}$ m/s	Prędkość średnia kwadratowa $\sqrt{\overline{v^2}}$ m/s	Średnia droga swobodna l 10^{-8} m	Liczba zderzeń w sekundzie n 10^9 s^{-1}	Średnica cząsteczki $2d$ 10^{-10} m
wodór	1692	1840	11,2	15,1	2,3
tlen	425	461	6,5	6,5	2,9
azot	454	493	6,0	7,5	3,1
hel	1204	1305	18,0	6,9	1,9
dwutlenek węgla	362	393	4,0	9,0	3,2
para wodna	566	615	4,0	14,1	2,6

16.11. Zjawiska transportu energii, masy i pędu

Najbliższe nasze rozważania będą dotyczyły zjawisk, którym nadamy wspólną nazwę zjawisk transportu (przenoszenia). Będą to zjawiska transportu energii (przewodnictwo cieplne), transportu masy (dyfuzja) i transportu pędu (lepkość). O przewodnictwie cieplnym i dyfuzji będziemy mówili po raz pierwszy, w związku z tym podamy podwójną charakterystykę tych zjawisk, a mianowicie charakterystykę makroskopową i charakterystykę cząsteczkową, opartą o teorię kinetyczno-molekularną. Makroskopowy opis zjawisk lepkości z § 9.8 uzupełnimy charakterystyką cząsteczkową.

Wszystkie trzy wymienione zjawiska mają tę wspólną cechę, że są to procesy kinetyczne, procesy ustalania się równowagi. Zakłócenie równowagi może być związane np. z wywołaniem gradientu temperatury (przewodnictwo cieplne), gradientu gęstości (ogólniej mówiąc gradientu stężeń) w spoczywającym ośrodku (dyfuzja) lub gradientu prędkości w poruszających się względem siebie warstwach cieczy lub gazów (lepkość). Stopniowe wyrównywanie się temperatur, gęstości (stężeń) lub prędkości warstw zachodzi w wymienionych procesach dzięki chaotycznemu ruchowi cieplnemu cząsteczek rozważanych ośrodków.



Rys. 16.9

16.11.1. *Przewodnictwo cieplne.* Przez *przewodnictwo cieplne* rozumiemy przenoszenie energii cieplnej wywołane istnieniem gradientu temperatury. Wyobraźmy sobie np. warstwę ciała o grubości dx (rys. 16.9), przez którą w kierunku dodatnim osi x przepływa energia cieplna. Niech badana warstwa będzie ograniczona dwoma przekrojami A i C, prostopadłymi do osi x, o powierzchni S i temperaturach odpowiednio T_1 i T_2. Z założenia dotyczącego kierunku przepływu energii cieplnej wynika, że $T_1 > T_2$. Gradient temperatury wynosi dT/dx i jest skierowany w stronę ujemnych wartości x. Między wspomnianymi przekrojami wybierzmy elementarną powierzchnię dS również prostopadłą do osi x.

W odniesieniu do przepływającej przez powierzchnię dS energii cieplnej można wprowadzić pojęcie strumienia energii $d\phi$ i gęstości strumienia energii $d\varphi$. Doświadczenia wykazują, że *strumień energii*, czyli ilość energii przechodzącej w jednostce czasu przez powierzchnię dS, ustawioną prostopadle do kierunku przepływu energii, jest proporcjonalny do pola powierzchni dS i do gradientu temperatury, a więc

$$\frac{dQ}{dt} = d\phi = -\lambda \frac{dT}{dx} dS \tag{16.9}$$

(znak minus we wzorze przypomina, że kierunek przepływu energii jest przeciwny do kierunku gradientu temperatury).

Odpowiednio, *gęstość strumienia energii* $d\varphi$ — czyli ilość energii przechodzącej w jednostce czasu przez jednostkę powierzchni ustawionej prostopadle do kierunku przepływu energii — równa się

$$d\varphi = \frac{d\phi}{dS} = -\lambda \frac{dT}{dx}. \tag{16.10}$$

Występujący w obu ostatnich wzorach współczynnik proporcjonalności λ wyraża się — jak łatwo sprawdzić — w J/m · s · K. Nosi nazwę współczynnika przewodzenia ciepła i wyraża liczbowo ilość energii cieplnej przenikającej w jednostce czasu przez jednostkę powierzchni ustawionej prostopadle do kierunku przepływu energii przy jednostkowym gradiencie temperatury. Innymi słowy — im większa jest wartość współczynnika, tym lepszym przewodnikiem ciepła jest dany materiał. Z tabeli 16.3 widać, że do najlepszych przewodników ciepła należą metale, do najgorszych — gazy. Wartość tego współczynnika dla określonego materiału można w dość szerokich granicach uważać za stałą, aczkolwiek w zasadzie wykazuje on powolny wzrost ze wzrostem temperatury.

Tabela 16.3

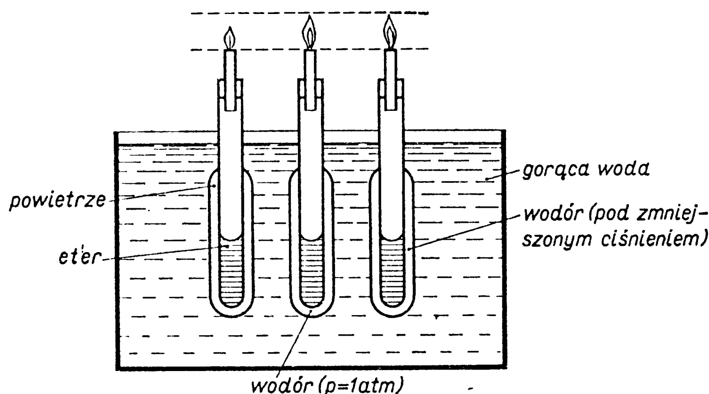
Współczynniki przewodzenia ciepła

Rodzaj substancji	λ J/m · s · K	t °C
chlor	$72 \cdot 10^{-4}$	0
dwutlenek węgla	$145 \cdot 10^{-4}$	0
tlen	$236 \cdot 10^{-4}$	0
powietrze	$240 \cdot 10^{-4}$	0
wodór	$1680 \cdot 10^{-4}$	0
woda	$60 \cdot 10^{-2}$	20
benzen	$15 \cdot 10^{-2}$	20
ołów	35	20
stal	46	20
miedź	380	20
srebro	420	20
azbest	$8 \cdot 10^{-3}$	
drewno	$8 \cdot 10^{-3}$	
beton	$8 \cdot 10^{-2}$	
korek	$16 \cdot 10^{-3}$	
lód	$16 \cdot 10^{-2}$	
ebonit	$17 \cdot 10^{-2}$	

Metale wykazują nie tylko najlepsze przewodnictwo cieplne, lecz także najlepsze przewodnictwo elektryczne. Do związku obu tych zjawisk, których mechanizm jest ten sam, powrócimy jeszcze w § 20.10 i 32.4.

Nawiązując do prostych doświadczeń zilustrujemy różnice przewodnictwa cieplnego metali, cieczy (konkretnie — wody) i gazów. Wiadomo, że pręta metalowego nie utrzymamy długo w ręce, jeśli drugi jego koniec będzie ogrzewany np. w płomieniu gazowym. Natomiast probówkę z wodą można będzie trzymać za dno nawet wtedy, gdy doprowadzimy wodę do wrzenia, ogrzewając ją w okolicy powierzchni swobodnej. (Podkreślamy różnicę w stosunku do normalnego ogrzewania wody — zwykle ogrzewamy wodę od dołu i wtedy głównym zjawiskiem decydującym o wymianie ciepła w całej objętości jest konwekcja. W opisanym ostatnio doświadczeniu ogrzewamy wodę od góry, a zatem wymiana ciepła może zachodzić wyłącznie przez przewodnictwo.)

Doświadczenie pokazane na rys. 16.10 dotyczy przewodnictwa gazów. W dużym naczyniu wypełnionym gorącą wodą zanurzone są trzy probówki zawierające jednakowe ilości eteru. Każda z probówek jest wtopiona w szersze naczynie. Pierwsze z tych naczyń zawiera powietrze o ciśnieniu normalnym, drugie — wodór również o ciśnieniu normalnym, trzecie — wodór o ciśnieniu zmniejszonym. Ciepło od gorącej wody jest przenoszone do eteru przede wszystkim dzięki przewodnictwu gazów. Im więcej ciepła otrzymuje eter w każdej jednostce czasu, tym gwałtowniej paruje. Parę eteru zapalamy u wylotu rurki wyprowadzonej przez korek z wnętrza probówki na zewnątrz. Wysokość płomienia świadczy o ilości ciepła przewodzonego przez wymienione gazy. Okazuje się, że wodór przewodzi ciepło lepiej niż powietrze. Wodór rozrzedzony przewodzi tak samo jak wodór o ciśnieniu normalnym. Jest to właściwość wszystkich gazów — *przewodnictwo cieplne gazów nie zależy od ciśnienia.*



Rys. 16.10

Przejdźmy obecnie do analizy zjawiska przewodnictwa cieplnego z punktu widzenia teorii kinetyczno-molekularnej, ograniczając się do rozważań związanych z przewodnictwem gazów.

Wiemy już, że różnice temperatur w różnych obszarach wiążą się z istnieniem różnic średnich energii kinetycznych cząsteczek w tych obszarach. Ale z biegiem czasu na skutek nieuniknionych zderzeń międzycząsteczkowych odbywa się stopniowe wyrównywanie się średnich E_k, a tym samym i temperatur.

W celu powiązania współczynnika λ z wielkościami mikroskopowymi, jak średnia długość drogi swobodnej l, średnia prędkość ruchu cieplnego cząsteczek itp., rozważmy energię kinetyczną przenoszoną przez powierzchnię dS prostopadłą do osi x (pomyślaną między warstwami o temperaturach T_1 i T_2) przez poruszające się cząsteczki gazu w kierunku od lewej do prawej i odwrotnie (rys. 16.9).

Można przyjąć, że każda z cząsteczek przenosi przez przekrój B tę ilość energii, którą uzyskała w czasie ostatniego zderzenia, które nastąpiło — średnio biorąc — w odległości równej średniej drodze swobodnej l od przekroju B. Ponieważ gradient temperatury w obszarze między A i C równa się dT/dx, więc różnica temperatur na od-

ległości l wynosi $(dT/dx)l$, z tym że po prawej stronie od B mamy spadek temperatury równy $(-dT/dx)l$, po lewej zaś wzrost równy $(dT/dx)l$.

Jeżeli założymy, że przekrojowi B odpowiada temperatura T, to energię kinetyczną cząsteczki dochodzącej do B od lewej strony można wyrazić wzorem

$$E'_k = mc_V\left(T + l\frac{dT}{dx}\right),$$

a od prawej — wzorem

$$E''_k = mc_V\left(T - l\frac{dT}{dx}\right),$$

gdzie m oznacza masę cząsteczki, c_V — ciepło właściwe w stałej objętości.

Wskutek ruchu tych cząsteczek wystąpi na przekroju B zmiana energii

$$E'_k - E''_k = 2mc_V\frac{dT}{dx}l.$$

Oznaczmy liczbę cząsteczek w jednostce objętości przez n, a ich prędkość średnią przez $\bar{v}$. Wtedy energia przeniesiona w jednostce czasu przez jednostkową powierzchnię wyciętą z B wyrazi się wzorem

$$d\varphi = \frac{2mc_V\frac{dT}{dx}ln\bar{v}}{6}.$$

Czynnik 6 w mianowniku pochodzi stąd, że prędkości cząsteczek gazowych pod względem kierunku są rozmieszczone zupełnie chaotycznie, a więc w kierunku dodatnim osi x porusza się $n/6$ cząsteczek.

$$d\varphi = \tfrac{1}{3}mn\bar{v}c_V l\frac{dT}{dx}. \tag{16.11}$$

Iloczyn mn (masy m pojedynczej cząsteczki i liczby n cząsteczek w jednostce objętości) jest gęstością ϱ gazu. A zatem

$$d\varphi = \tfrac{1}{3}\varrho\bar{v}c_V l\frac{dT}{dx}. \tag{16.12}$$

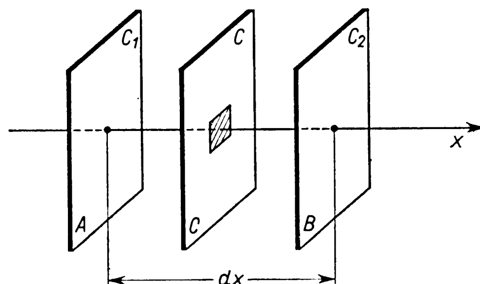
Z porównania otrzymanego wzoru z wyrażeniem (16.10) wynika powiązanie współczynnika przewodzenia ciepła λ z wielkościami mikroskopowymi, a mianowicie

$$\lambda = \tfrac{1}{3}\varrho\bar{v}c_V l. \tag{16.13}$$

Uwzględniając fakt, że średnia droga swobodna jest odwrotnie proporcjonalna do ciśnienia gazu, a gęstość jest do ciśnienia wprost proporcjonalna, znajdziemy, że iloczyn ϱl nie zależy od ciśnienia. Ponieważ $\bar{v}$ jest od ciśnienia niezależne, więc λ gazu od ciśnienia nie zależy. Natomiast średnia prędkość $\bar{v}$ wzrasta z temperaturą, a zatem taką samą zależność wykazuje współczynnik przewodzenia ciepła λ gazu.

16.11.2. *Dyfuzja.* Jak już mówiliśmy, warunkiem wystąpienia procesu dyfuzji jest istnienie różnicy stężeń w różnych punktach obszaru zajmowanego przez dyfundujące

ciała. Należy tu od razu podkreślić, że termin stężenie c może mieć różne znaczenia, może oznaczać np. stężenie wagowe (masa przypadająca na jednostkę objętości), stężenie cząsteczkowe (liczba cząsteczek przypadająca na jednostkę objętości) lub stężenie molowe (liczba moli przypadająca na jednostkę objętości). Niezależnie od tego, jaki rodzaj stężenia mamy na myśli mówiąc o istnieniu gradientu stężenia w danym obszarze, dyfuzja zawsze sprowadza się do transportu masy z obszaru o większym do obszaru o mniejszym stężeniu.



Rys. 16.11

Wprowadźmy założenie upraszczające, że stężenie c jest tylko funkcją x i np. maleje w kierunku dodatniej osi x. Transport masy odbywa się w kierunku malejącego stężenia, a zatem, np. od warstwy A (rys. 16.11) o stężeniu większym c_1 do warstwy B o stężeniu mniejszym c_2. Obie warstwy są odległe od siebie o dx. Gradient stężenia między warstwami równa się dc/dx (skierowany jest w kierunku ujemnej osi x).

Weźmy pod uwagę elementarną powierzchnię dS między warstwami A i B, prostopadłą do osi x. Strumień dyfuzji $d\phi_d$ przez tę powierzchnię wyraża liczbowo masę przenikającą w jednostce czasu w kierunku dodatniej osi x przez tę powierzchnię. Z badań doświadczalnych wynika, że *strumień dyfuzji* jest proporcjonalny do gradientu stężenia i do pola powierzchni dS, czyli

$$d\phi_d = -D\frac{dc}{dx}dS,$$

gdzie D oznacza współczynnik proporcjonalności, zwany *współczynnikiem dyfuzji*. Znak minus we wzorze przypomina, że kierunek strumienia dyfuzji jest przeciwny względem kierunku gradientu stężenia.

Od strumienia dyfuzji można przejść do gęstości tego strumienia. *Gęstość strumienia dyfuzji* $d\varphi_d$ wyraża liczbowo masę przenikającą w jednostce czasu przez jednostkę powierzchni ustawionej prostopadle do kierunku dyfuzji (w naszym przypadku do kierunku osi x):

$$d\varphi_d = -D\frac{dc}{dx}. \qquad (16.14)$$

Sens fizyczny współczynnika dyfuzji D znajdujemy uwzględniając we wzorze (16.14) jednostkowy gradient stężenia ($dc/dx = 1$). *Współczynnik dyfuzji* wyraża liczbowo gę-

stość strumienia dyfuzji w przypadku jednostkowego gradientu stężenia. Jednostką współczynnika dyfuzji — jak widać ze wzoru $D = -\frac{d\varphi_d}{dc/dx}$ — w układzie SI jest m^2/s. Ważną jest rzeczą, że współczynnik dyfuzji nie zależy od sposobu zdefiniowania stężenia (jeśli bowiem stężenie jest zdefiniowane wagowo lub cząsteczkowo, to i gęstość strumienia odpowiednio odnosi się do masy lub liczby cząsteczek dyfundujących przez jednostkę powierzchni w jednostce czasu).

Współczynnik dyfuzji zależy od rodzaju ciała dyfundującego i od ośrodka, do którego ono dyfunduje, jak również od warunków fizycznych charakteryzujących ośrodek, jak gęstość i temperatura.

Rząd wielkości współczynników dyfuzji gazów w normalnych warunkach waha się w granicach 10^{-5} m^2/s–10^{-4} m^2/s (tab. 16.4). Dla cieczy jest on 10^4–10^5 razy mniejszy. W ciałach stałych proces dyfuzji — również stwierdzony doświadczalnie — jest jeszcze znacznie (około 10^6 razy) powolniejszy.

Tabela 16.4

Współczynniki dyfuzji gazów

(temperatura 0°C; ciśnienie normalne)

Rodzaj gazu	D 10^{-6} m^2/s
wodór — powietrze	64
wodór — tlen	70
tlen — powietrze	17
CO_2 — powietrze	14
tlen — azot	17

Warto może podkreślić, że w przypadku stykających się ze sobą różnych cieczy nie zawsze występuje dyfuzja: istnieją *ciecze* tzw. *nie mieszające się*, jak np. rtęć i woda, w których praktycznie biorąc dyfuzji się nie obserwuje. Innym rodzajem zestawu cieczy są *ciecze idealnie mieszające się*, tzn. mieszające się w każdym dowolnym stosunku (np. woda i alkohol, woda i eter). Są też ciecze mieszające się tylko do pewnego określonego stopnia. Po osiągnięciu tego granicznego stosunku w mieszaninie badanych cieczy nadmiar którejkolwiek z nich nie ulega już dalszemu wymieszaniu: procesu dyfuzji już się nie obserwuje.

W przypadku stykających się ze sobą gazów dyfuzja występuje zawsze, jeśli tylko gazy nie wchodzą ze sobą w gwałtowną reakcję chemiczną.

Przeanalizujmy z kolei zjawisko dyfuzji w gazach z punktu widzenia teorii kinetyczno-molekularnej. Gdyby cząsteczki gazów poruszały się bez zderzeń, proces dyfuzji zachodziłby bardzo szybko, tego jednak doświadczenia nie wykazują. Wskutek zderzeń międzycząsteczkowych każda cząsteczka gazu, mimo że przebiega w czasie 1 s drogę rzędu kilkuset metrów (por. wartości prędkości średnich cząsteczek gazu — tab. 16.2), to jednak ostatecznie w tym czasie mało się oddala od swego położenia początkowego

(por. rys. 16.2). Im większa jest liczba zderzeń na sekundę, czyli im mniejsza jest średnia droga swobodna, tym wolniej przebiega proces dyfuzji.

Rozumując podobnie jak w przypadku przewodnictwa cieplnego dojdziemy do nowego „mikroskopowego" wyrażenia na gęstość strumienia dyfuzji. Rys. 16.11 może nam oddać usługi i w tym przypadku. Niech powierzchnia A odpowiada warstwie gazu o stężeniu c_1, powierzchnia B — warstwie o nieco mniejszym stężeniu c_2. Wobec wzajemnej odległości między warstwami równej dx gradient stężenia wynosi dc/dx. Cząsteczki gazu w swoim chaotycznym ruchu cieplnym przechodzą przez powierzchnię dS w kierunku od prawej strony ku lewej i odwrotnie. Stosując te same przesłanki, co przy badaniu przewodnictwa cieplnego, znajdziemy nadwyżkę masy przenikającej przez jednostkową powierzchnię prostopadłą do osi x w jednostce czasu, czyli gęstość strumienia dyfuzji

$$d\varphi_d = \tfrac{1}{3}\bar{v}l\,\frac{dc}{dx}.$$

Porównując ten wzór z (16.14) otrzymujemy

$$D = \tfrac{1}{3}\bar{v}l. \tag{16.15}$$

A zatem współczynnik dyfuzji jest proporcjonalny do średniej prędkości ruchu cieplnego cząsteczek gazu i do średniej drogi swobodnej. Za tym się kryją oczywiście zależności od takich wielkości makroskopowych, jak ciśnienie i temperatura, bowiem średnia droga swobodna jest odwrotnie proporcjonalna do ciśnienia gazu, a średnia prędkość ruchu cieplnego rośnie wraz z temperaturą. Tak więc z rozważań teoretycznych wynika odwrotna proporcjonalność współczynnika dyfuzji D do ciśnienia gazu i wzrost jego wartości ze wzrostem temperatury. Obie te zależności potwierdza doświadczenie.

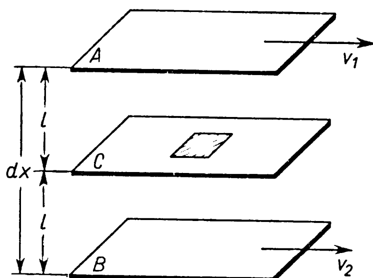
Trzeba podkreślić, że przy wyprowadzaniu wzoru (16.15) wprowadziliśmy uproszczenie polegające na przypisaniu cząsteczkom gazu dyfundującego i gazu stanowiącego środowisko tych samych wartości $\bar{v}$ i l. Można wykazać, że w przypadku dużych różnic wymienionych wartości dla obu gazów należy we wzorze (16.15) uwzględniać większą prędkość średnią i mniejszą średnią drogę swobodną.

16.11.3. *Lepkość gazów.* Makroskopowemu przedstawieniu zjawiska lepkości poświęcony był § 9.10. Obecnie zajmiemy się tylko rozpatrzeniem zjawiska lepkości gazów z punktu widzenia teorii kinetyczno-molekularnej.

Niech laminarny przepływ gazu odbywa się w kierunku poziomym. W tych warunkach prędkość przepływu gazu w poszczególnych warstwach poziomych jest stała, zmienia się natomiast przy przejściu od warstwy do warstwy. Załóżmy np., że maleje ona w miarę przesuwania się ku dołowi. Jeśli zatem przyjmiemy, że tym razem oś x jest skierowana pionowo, to prędkość przepływu gazu będzie tylko funkcją x.

Weźmy pod uwagę dwie warstwy poziome A i B (rys. 16.12), odległe od siebie o dx, którym odpowiadają prędkości przepływu gazu v_1 i nieco mniejsza v_2. Między warstwami istnieje ciągły spadek prędkości, który powoduje wystąpienie gradientu prędkości przepływu dv/dx.

Wybierzmy pomocniczą płaszczyznę C, leżącą między A i B, równoległą do nich. Cząsteczki gazu nad tą płaszczyzną mają na ogół prędkości, a zatem i pędy, większe niż cząsteczki znajdujące się poniżej. Dzięki chaotycznemu ruchowi cieplnemu w każdej sekundzie przez jednostkową powierzchnię wyciętą z warstwy C przechodzi pewna liczba cząsteczek gazu z dołu do góry i tyleż z góry na dół. Wskutek tego ogólny zasób pędu warstwy górnej maleje, a dolnej — rośnie. Ale wiemy, że zgodnie z drugą zasadą dynamiki zmiana pędu danego ciała, zachodząca w jednostce czasu, jest miarą działającej na to ciało siły. A zatem wzrost pędu masy gazu znajdującej się poniżej C wynika z działania na tę masę siły o kierunku zgodnym z v_2, natomiast zmniejszenie pędu masy gazu znajdującego się nad warstwą C wiąże się z siłą działającą w kierunku przeciwnym do v_1. A więc pierwsze z wymienionych sił przyspieszają masy gazu poruszające się wolniej, a drugie — zwalniają ruch mas gazu poruszających się szybciej. Te siły już znamy — są to siły tarcia wewnętrznego, czyli siły lepkości.



Rys. 16.12

Wytnijmy z warstwy C jednostkową powierzchnię i obliczmy na niej *gęstość strumienia pędu*. Aczkolwiek przenoszenie pędu jest związane z ruchem cieplnym cząsteczek, to jednak o bilansie pędu decydują nie pędy „cieplne" (przenoszone średnio jednakowo w obu kierunkach przez C), lecz pędy związane z makroskopowym ruchem gazu. Biorąc pod uwagę prędkości przepływu gazu w odległości średniej drogi swobodnej l po obu stronach C i uwzględniając fakt, że masy przenoszone w obu kierunkach przez jednostkową powierzchnię w jednostce czasu są jednakowe i równe $\frac{1}{6}\varrho\bar{v}$, znajdziemy, że gęstość strumienia pędu równa się

$$m\left(v+l\frac{dv}{dx}\right)-m\left(v-l\frac{dv}{dx}\right)=\frac{1}{6}\varrho\bar{v}\cdot 2l\frac{dv}{dx}=\frac{1}{3}\varrho\bar{v}l\frac{dv}{dx}.$$

Jest to równocześnie miara siły lepkości działającej na jednostkę powierzchni poruszającej się warstwy gazu. Można ją zatem przyrównać do odniesionej do tych samych warunków siły lepkości wyrażonej „makroskopowym" wzorem Newtona (por. wzór (9.16)). A zatem

$$\frac{1}{3}\varrho\bar{v}l\frac{dv}{dx}=\eta\frac{dv}{dx},$$

skąd

$$\eta=\frac{1}{3}\varrho\bar{v}l. \tag{16.6}$$

Jak to już podkreślaliśmy, iloczyn ϱl (podobnie jak i $\bar{v}$) nie zależy od ciśnienia gazu, a więc ciśnienie w nieograniczonym ośrodku gazowym nie ma wpływu na współczynnik lepkości gazu. Nie dotyczy to jednak ani gazów bardzo rozrzedzonych w zbiornikach zamkniętych, ani gazów bardzo zgęszczonych. W tym ostatnim przypadku trudno jest w ogóle mówić o średniej drodze swobodnej.

Współczynnik lepkości gazów zależy od temperatury, gdyż $\bar{v}$ jest funkcją temperatury. Ze wzrostem temperatury rośnie $\bar{v}$ i rośnie η (jest to zależność odwrotna niż w przypadku cieczy).

Kończąc omawianie zjawisk przewodnictwa cieplnego, dyfuzji i lepkości raz jeszcze podkreślamy wspólne cechy tych zjawisk. W każdym z nich mamy do czynienia z transportem pewnej wielkości, a mianowicie odpowiednio z transportem energii cieplnej, masy i pędu. Przebieg każdego z tych procesów w środowisku gazowym jest zależny od długości średniej drogi swobodnej. W związku z tym metody wyznaczania średniej drogi swobodnej oparte są na badaniach doświadczalnych tych procesów, a w szczególności lepkości.

16.12. Ciała krystaliczne i bezpostaciowe. Podstawowe wiadomości z krystalografii

W paragrafie niniejszym i następnych rozważać będziemy właściwości ciał stałych. Dla uniknięcia nieporozumień rozpoczniemy od rozróżnienia ciał krystalicznych (stałych) i bezpostaciowych (przechłodzone ciecze).

Ciała krystaliczne modelowe, tzw. doskonałe kryształy, charakteryzuje nieskończona, regularna powtarzalność niezmiennych elementów strukturalnych. Ciała bezpostaciowe tej regularności nie wykazują: atomy lub cząsteczki rozmieszczone są w nich w sposób chaotyczny. Jako przykłady ciał krystalicznych wymienić można metale, lód, sól kamienną itp. Wśród ciał bezpostaciowych można wymienić szkło, smołę, żywice itp. Jak już wspomnieliśmy, ciała bezpostaciowe traktuje się jako przechłodzone ciecze, tj. ciecze oziębione bez zestalenia poniżej temperatury krzepnięcia (por. § 17.3). Ciecze takie charakteryzuje bardzo duży współczynnik lepkości.

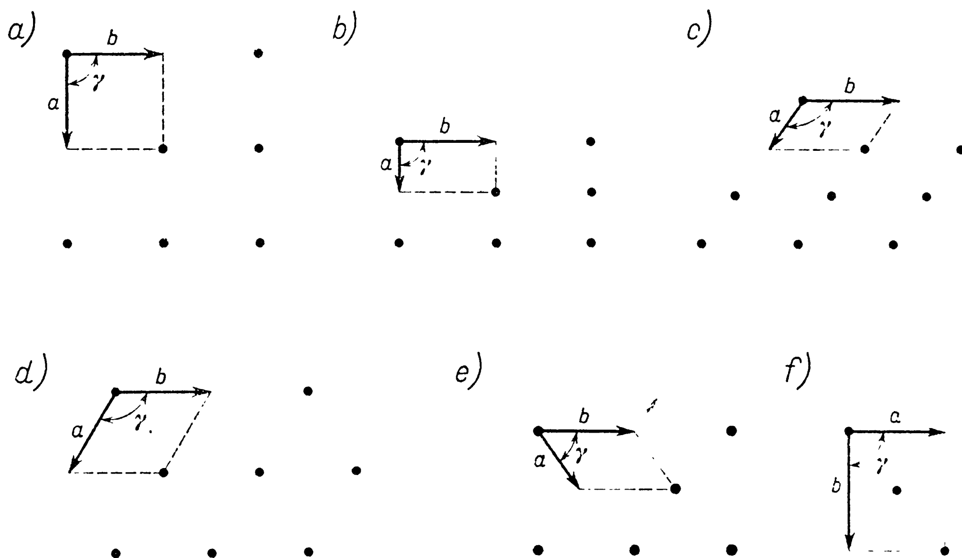
Jedna z podstawowych różnic właściwości ciał krystalicznych i bezpostaciowych sprowadza się do anizotropii kryształów i izotropii ciał bezpostaciowych. Anizotropia kryształów polega na tym, że w różnych kierunkach ciała te wykazują różne właściwości. Jako przykład mogą służyć znane już nam: anizotropia właściwości mechanicznych (różne wartości modułów sprężystości dla różnych kierunków) i anizotropia właściwości cieplnych (różne wartości przewodności dla różnych kierunków) oraz anizotropia optyczna i elektryczna, o których jeszcze będziemy mówili. Jednym z efektów anizotropii mechanicznej jest także występowanie w kryształach określonych płaszczyzn łupliwości.

W niektórych ciałach o budowie krystalicznej anizotropia jest słabo zaznaczona. Przy bliższych badaniach często się okazuje, że ciała takie mają *budowę* tzw. *polikrystaliczną* (w odróżnieniu od *budowy monokrystalicznej*), tzn. składają się z wielkiej liczby

chaotycznie rozmieszczonych małych kryształków. To chaotyczne ułożenie decyduje właśnie o zaniku makroskopowego efektu anizotropii.

Powróćmy do uwag dotyczących budowy ciał monokrystalicznych. Dzięki powtarzalności elementów strukturalnych opartej na różnych zasadach (przesunięcia, obroty itp.) kryształy wykazują określone prawidłowości w budowie geometrycznej, jak również określone cechy symetrii. Te prawidłowości w budowie geometrycznej i elementy symetrii wykorzystano m. in. do wprowadzenia pewnej klasyfikacji kryształów.

Dzięki regularnemu rozmieszczeniu w przestrzeni elementów struktury kryształu można wprowadzić pojęcie *sieci krystalicznej*, zdefiniowanej jako równoległe uszeregowanie punktów, zwanych *węzłami*, mających tę własność, że otoczenie każdego z nich jest jednakowe. Opis struktury kryształu wymaga znajomości geometrii sieci (tzn. odległości międzywęzłowych i kątów między nimi), jak również przyporządkowania węzłom określonego atomu, jonu lub też zespołu atomów ściśle określonego pod względem składu i rozmieszczenia przestrzennego składników.



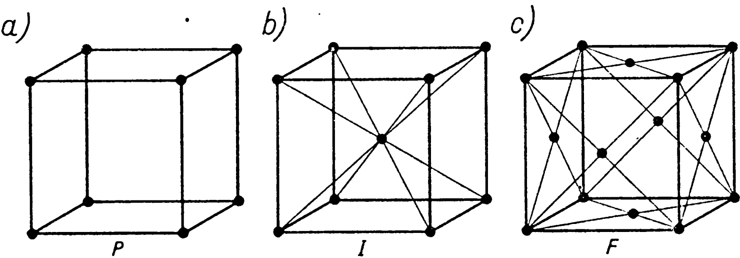
Rys. 16.13

Różne rodzaje sieci zilustrujemy graficznie ograniczając się — dla prostoty ujęcia — do kryształów dwuwymiarowych (rys. 16.13). Wielkości a i b oznaczają odległości międzywęzłowe, a kąt γ — kąt między nimi. Kolejne rysunki 16.13a, b, c i d przedstawiają odpowiednio sieć kwadratową, prostokątną, skośnokątną i heksagonalną. Ich charakterystyki geometryczne podane są w załączonym zestawieniu.

Na każdym z wymienionych rysunków na odcinkach a i b zbudowano równoległoboki, uzyskując w ten sposób tzw. komórkę prostą kryształu, czyli taką odmianę komórki elementarnej kryształu, w której węzły sieci znajdują się wyłącznie w wierzchołkach komórki.

Nazwa sieci dwuwymiarowego kryształu	Charakterystyka geometryczna komórki prostej kryształu
kwadratowa	$a = b \quad \gamma = 90°$
prostokątna	$a \neq b \quad \gamma = 90°$
skośnokątna	$a \neq b \quad \gamma \neq 90°$
heksagonalna	$a = b \quad \gamma = 120°$
prostokątna scentrowana	$a \neq b \quad \gamma \neq 90°$ komórka prosta $a \neq b \quad \gamma = 90°$ komórka elementarna z jednym węzłem w środku

Rysunki 16.13e i f przedstawiają ostatni — piąty — rodzaj sieci dwuwymiarowego kryształu, a mianowicie sieć prostokątną scentrowaną. W stosunku do sieci prostokątnej prostej, przedstawionej na rys. 16.13b, występuje tu dodatkowy węzeł na przecięciu przekątnych prostokąta. W tym przypadku można wyróżnić komórkę prostą w postaci równoległoboku zaznaczonego na rys. 16.13e o charakterystyce $a \neq b$, $\gamma \neq 90°$ oraz komórkę elementarną (z jednym węzłem w środku komórki) w postaci prostokąta zaznaczonego na rys. 16.13f o charakterystyce $a \neq b$ i $\gamma = 90°$. Warto podkreślić jednakową charakterystykę komórki elementarnej sieci prostokątnej scentrowanej i komórki prostej sieci prostokątnej niescentrowanej.

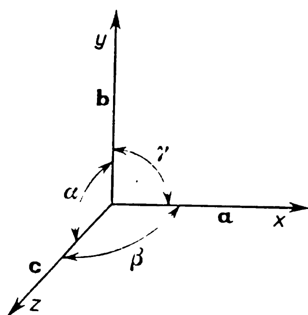


Rys. 16.14

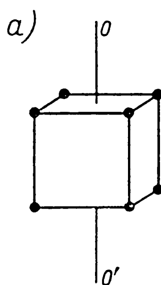
W odniesieniu do trójwymiarowych sieci krystalicznych sprawa się oczywiście komplikuje. Dla przykładu weźmy pod uwagę przestrzenny odpowiednik sieci kwadratowej (rys. 16.14). Jak widać, istnieją tu już trzy możliwości: obok sieci regularnej prostej (P) istnieje sieć regularna scentrowana przestrzennie (I) (rys. 16.14b) oraz sieć regularna scentrowana powierzchniowo (F) (rys. 16.14c). Zamiast pięciu możliwych dwuwymiarowych sieci krystalicznych otrzymuje się ogółem 14 sieci przestrzennych. Każdej z tych czternastu sieci przestrzennych odpowiada pewna komórka elementarna stanowiąca już twór przestrzenny, a nie płaski, jak to miało miejsce w przypadku sieci dwuwymiarowych. Komórkami elementarnymi mogą być np. sześciany, graniastosłupy proste o podstawach kwadratowych lub rombowych, graniastosłupy skośne o podstawie rombowej itp.

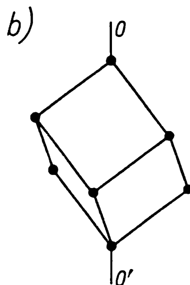
W zależności od odległości międzywęzłowych a, b, c komórek elementarnych oraz kątów α, β, γ, jakie tworzą ze sobą wektory **a**, **b**, **c** (rys. 16.15), wprowadzono podział wymienionych czternastu sieci na siedem układów. Są to *układy*: *trójskośny*, *jednoskośny*, *rombowy*, *tetragonalny*, *regularny*, *trygonalny* i *heksagonalny*. Podamy przykładowo charakterystyki dwóch z nich, a mianowicie:

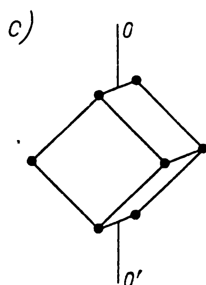
w układzie regularnym $a = b = c$ $\alpha = \beta = \gamma = 90°$,
w układzie rombowym $a \neq b \neq c$ $\alpha = \beta = \gamma = 90°$.



Rys. 16.15







Rys. 16.16

Każdemu z wymienionych siedmiu układów odpowiada określona liczba elementów symetrii, wśród których m.in. występują jedno-, dwu-, trzy-, cztero- i sześciokrotne osie obrotu oraz płaszczyzny zwierciedlanych odbić. Rys. 16.16a, b, c przedstawiają odpowiednio przykłady cztero-, trzy- i dwukrotnej osi symetrii w układzie regularnym. *Krotność osi* wskazuje, ile razy podczas obrotu o 360° kryształ zajmie identyczne położenie w przestrzeni.

16.13. Klasyfikacja wiązań krystalicznych

Od rozważań związanych z geometrią kryształów przejdźmy do omówienia klasyfikacji z punktu widzenia charakteru sił zapewniających trwałość budowy krystalicznej. Zagadnienie jest skomplikowane z tego względu, że o spójności kryształu rzeczywistego decydują na ogół różne rodzaje sił wiązania, chociaż niejednokrotnie jeden rodzaj sił ma wybitną przewagę nad innymi. Z tego względu ograniczymy się do wskazania *kryształów modelowych*, tzn. idealnych typów ciał stałych o ściśle określonym rodzaju sił wiązania.

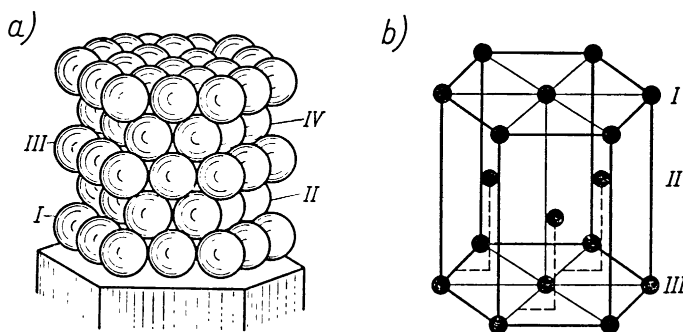
Do idealnych typów ciał stałych zaliczamy: 1) kryształy metaliczne, 2) kryształy jonowe, 3) kryształy walencyjne, 4) kryształy molekularne.

Stosując na razie tylko podstawowe wiadomości ze szkoły średniej podamy możliwie proste charakterystyki sił wiązania, występujących w poszczególnych typach idealnych kryształów.

Zaczniemy nasz przegląd od *kryształów metalicznych*. Jak wiadomo, metale odznaczają się dobrym przewodnictwem cieplnym i elektrycznym, wynikającym z istnienia w metalu elektronów swobodnych, zwanych też elektronami przewodnictwa. Można

więc wyobrazić sobie kryształ metaliczny w postaci uporządkowanego przestrzennie zbioru dodatnich jonów metalu, umieszczonych w środowisku wypełnionym elektronami przewodnictwa. Siły elektrycznego przyciągania między dodatnimi jonami i ich elektronowym ujemnym otoczeniem mają przewagę nad odpychaniem kulombowskim jonów dodatnich i decydują o siłach wiązania w idealnych kryształach metalicznych. Siły wiązania w tym przypadku nie są duże i w związku z tym tzw. *energia wiązania*, tzn. energia potrzebna do rozdzielenia ciała stałego na oddzielne atomy, cząsteczki lub jony, również nie jest zbyt duża. Tak np. dla metalicznego sodu wynosi ona w temperaturze pokojowej około 1,1 elektronowoltów* na cząsteczkę, co równa się $6{,}6 \cdot 10^{17}$ MeV/mol.

Wspomniane siły nie mają uprzywilejowanego kierunku w przestrzeni, raczej można mówić o ich równomiernym, przestrzennym rozkładzie. Z tym się wiąże tendencja metali do krystalizowania w takich układach, w których zgęszczenie atomów jest możliwie duże. Układami krystalograficznymi o najgęstszym „upakowaniu" elementów są układy heksagonalny i regularny scentrowany powierzchniowo. Jeśli przyjmiemy, że jony w krysztale metalicznym mają postać sztywnych kul o jednakowych średnicach, to stopień wypełnienia objętości w układzie heksagonalnym będzie sięgał 74%**. Reszta objętości — to nieuniknione wolne przestrzenie między jonami.



Rys. 16.17

Dla przykładu omówimy rozkład jonów w układzie heksagonalnym o najgęstszym upakowaniu. Można go modelowo przedstawić w postaci równoległych warstw złożonych z jednakowych kul. W warstwie pierwszej wyróżnić można fragment złożony z siedmiu kul, a mianowicie z kuli środkowej i sześciu stykających się z nią analogicznych kul (rys. 16.17). Nad tym fragmentem, w warstwie drugiej znajdują się trzy kule stykające się z sobą i odpowiednio z trzema kulami z warstwy niższej. Kule trzeciej warstwy przypadają dokładnie nad kulami warstwy pierwszej, a kule warstwy czwartej — nad kulami warstwy drugiej itd.

---

* 1 elektronowolt = $1{,}602 \cdot 10^{-19}$ J.

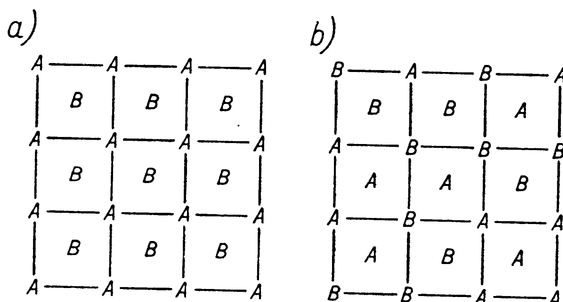
** Dla porównania podamy, że stopień wypełnienia objętości kryształu przez sztywne kule w układzie regularnym prostym wynosi ok. 52%.

W układzie heksagonalnym krystalizują: cynk, kobalt, magnez itd. W układzie regularnym scentrowanym powierzchniowo krystalizują m.in. aluminium, miedź, nikiel, ołów, platyna, srebro, złoto.

Gdy mowa o kryształach metalicznych, trudno nie wspomnieć o stopach metali. Rozpatrzmy przykłady *stopów podwójnych*, tzn. roztworów stałych dwóch metali.

Badania stopów różnych metali wykazały różnorodność odmian. Tak np. dwa metale o jednakowej wartościowości, krystalizujące w tym samym układzie i mające atomy o jednakowej (w granicach do 15%) średnicy, mogą mieszać się ze sobą w każdym dowolnym stosunku. Rozmieszczenie obu metali w węzłach sieci jest przypadkowe, z tym jednak, że raczej sąsiadują ze sobą atomy różnego rodzaju. Jako przykłady metali tworzących stopy niezależnie od ilości użytych składników można wymienić Ag i Au lub Cu i Ni.

Częściej jednak występują zestawienia metali o ograniczonych możliwościach tworzenia stopów. Do takich należą stopy np. miedź-cynk, miedź-kadm, magnez-antymon itd. Atomy obu metali wchodzących w skład stopu rozlokowują się w węzłach sieci w sposób przypadkowy lub uporządkowany (rys. 16.18).



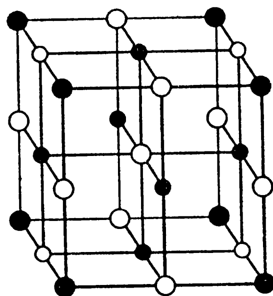
Rys. 16.18

Stopień uporządkowania wpływający na wiele właściwości stopu maleje ze wzrostem temperatury aż do tzw. *temperatury przejścia*, przy której znika uporządkowanie. *Zahartowanie stopu* (przejście gwałtowne od temperatury wysokiej do niskiej) może spowodować „zamrożenie" nieuporządkowania. Można też zniszczyć uporządkowanie w niskiej temperaturze przez napromieniowanie szybkimi cząstkami.

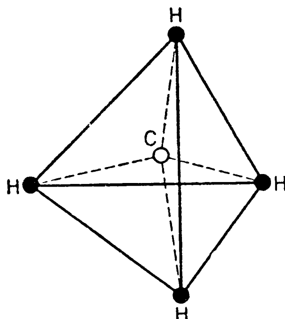
Innym rodzajem stopów są tzw. *stopy międzywęzłowe*. Tego rodzaju stopy mogą się wytworzyć wtedy, gdy atomy drugiego składnika są tak małe, że mieszczą się w wolnych przestrzeniach między atomami pierwszego składnika. Takie stosunki są charakterystyczne dla stali: atomy węgla o małych rozmiarach rozlokowują się w wolnych przestrzeniach między atomami żelaza.

Drugi rodzaj kryształów — to *kryształy jonowe*, charakteryzujące się tym, że występuje w nich częściowe przekazanie elektronów od atomów jednego rodzaju do atomów drugiego rodzaju. Towarzyszy temu powstawanie jonów dodatnich i ujemnych. Taką sytuację mamy w kryształach NaCl (rys. 16.19). Jednowartościowy sód przekazuje swój jedyny elektron wartościowości atomowi chloru redukując w ten sposób konfigurację

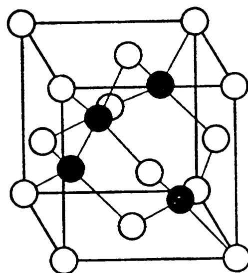
swych elektronów do wypełnionych powłok i zamykając równocześnie w ten sposób rozbudowę zewnętrznej powłoki elektronowej chloru (rozdz. 30 i 32). W ten sposób powstają jony Na^+ (czarne kółka na rysunku) i jony Cl^- (jasne kółka) rozmieszczone regularnie w przestrzeni w ten sposób, że każdy jon Na ma w swym najbliższym otoczeniu sześć jonów chloru, a każdy jon chloru — sześć jonów sodu. Dzięki mniejszym odległościom między jonami o znakach przeciwnych siły przyciągania elektrycznego jonów różnoimiennych przeważają nad siłami odpychania elektrycznego jonów jednoimiennych, dając ostatecznie duże siły wiążącego wypadkowego oddziaływania kulombowskiego między jonami.



Rys. 16.19



Rys. 16.20



Rys. 16.21

Ogólnie biorąc, dla kryształów tego typu charakterystyczna jest wysoka wartość energii wiązania (np. dla NaCl — około 7 razy większa niż dla metalicznego Na). Poza tym — wobec dużych sił wiążących — kryształy te wykazują dużą wytrzymałość na rozerwanie. Odznaczają się też małą ściśliwością, dużą twardością oraz wysoką temperaturą topnienia.

Do tej grupy kryształów należą liczne sole, jak np. NaCl, KCl, LiF, KBr, MgO, CaO itp.

Następny rodzaj kryształów — to *kryształy walencyjne*, złożone czasem nawet z atomów jednego pierwiastka (np. kryształy diamentu, krzemu, germanu). Wiązanie walencyjne między atomami jest wiązaniem elektronowym: przyjmuje się, że dwa atomy tworzące wiązanie przekazują po jednym elektronie do obszaru istniejącego między nimi.

Elektrony walencyjne obu atomów stają się jakby wspólne. Gęstość elektronów w obszarze między atomami staje się duża. W tych warunkach powstają bardzo duże ukierunkowane oddziaływania walencyjne. Naturę tego rodzaju wiązań można wytłumaczyć na podstawie mechaniki kwantowej.

Typowym przykładem wiązania walencyjnego jest wiązanie węgla: w licznych związkach węgla każdy jego atom tworzy cztery wiązania (np. w CH_4 lub w diamencie) tak rozmieszczone przestrzennie, jak przedstawia rys. 16.20. Uwzględniając powiązanie z czworościanem (rys. 16.20) mówi się często o czworościennym (tetraedrycznym) wiązaniu walencyjnym. Kąt między poszczególnymi wiązaniami wynosi około 109°. Do omówienia tego typu wiązania i ściśle z nim związanej struktury diamentu (rys. 16.21) wrócimy w rozdz. 32. Tutaj tylko jeszcze uzupełnimy ogólną charakterystykę kryształów walencyjnych wspominając o ich dużej energii wiązania i bardzo wysokiej twardości.

Na zakończenie tego przeglądu podamy jeszcze nieco danych o *kryształach molekularnych.* Siły wiązania występujące w tym przypadku są siłami oddziaływań międzycząsteczkowych (siły van der Waalsa). Jak wynika z kwantowomechanicznej teorii budowy cząsteczek, siły te mają charakter złożony i obejmują oddziaływania elektryczne różnego rodzaju (kulombowskie, dipolowe*), oddziaływania magnetyczne i siły tzw. kwantowe. Niektóre z tych sił są siłami przyciągania, inne — siłami odpychania. Wypadkowa siła zależy od odległości międzycząsteczkowych.

Siły wiązania van der Waalsa w kryształach molekularnych są stosunkowo słabe i idą w parze z niskimi temperaturami topnienia tych kryształów.

Liczne związki organiczne krystalizują właśnie w postaci kryształów molekularnych.

Może nasunąć się pytanie, co właściwie decyduje o ściśle określonym uporządkowanym układzie elementów struktury dowolnego kryształu. Układ elementów struktury jest taki, że zapewnia równowagę działających na każdy element sił przyciągających i odpychających ze strony otoczenia, a równocześnie zapewnia każdemu elementowi minimum energii potencjalnej. Każdy następny element „dobudowuje się" do kryształu przy spełnieniu tych samych warunków.

Każde działanie zewnętrzne, zmieniające wzajemne rozmieszczenie elementów kryształu, pociąga za sobą wzrost energii potencjalnej i wystąpienie sił wewnętrznych (które nazwaliśmy siłami sprężystymi). Siły te usiłują sprowadzić jony, atomy lub cząsteczki w kryształach do ich położeń równowagi, odpowiadających stanowi minimum energii potencjalnej. Źródłem tych sił sprężystych są oddziaływania elektryczne.

16.14. Defekty wewnątrz kryształów. Dyslokacje

Dotychczasowe nasze rozważania dotyczyły kryształów doskonałych o niezakłóconej powtarzalności rozmieszczenia przestrzennego ich elementów. W kryształach rzeczy-

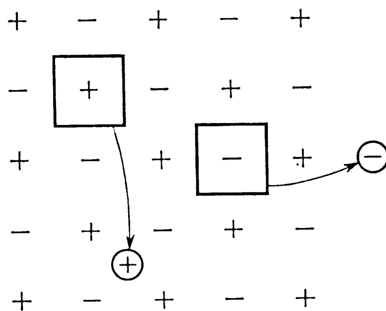
---

* *Dipolem elektrycznym* nazywamy układ dwóch ładunków o jednakowej wartości, lecz przeciwnych znakach, rozsuniętych na pewną odległość. Wytwarzanie się dipoli elektrycznych w wyniku zjawiska polaryzacji omówimy szerzej w rozdz. 19.

wistych istnieją różnego typu defekty. Badania z dziedziny fizyki ciała stałego wykazały, że natura tych defektów i ich liczba ma bardzo istotny wpływ na różne właściwości ciał stałych. Tak np. obce domieszki w kryształach decydują w dużej mierze o przewodności elektrycznej półprzewodników (rozdz. 32), o właściwościach optycznych (luminescencji) wielu kryształów, o właściwościach sprężystych itp. Nic więc dziwnego, że defektom w kryształach poświęcono bardzo wiele prac badawczych. Omówimy naturę kilku rodzajów defektów.

Najprostszym defektem budowy kryształów, zwanym *defektem Schottky'ego*, jest luka (dziura) w sieci — np. brak jonu dodatniego lub ujemnego w węźle sieci NaCl.

Innym rodzajem defektu sieci jest *defekt Frenkla*, polegający na przesunięciu elementu sieci z węzła do położenia międzywęzłowego (a więc również rodzaj luki w sieci). Jak wykazują badania, ten rodzaj defektu jest bardzo rozpowszechniony w kryształach halogenków srebra. Defekty Schottky'ego i Frenkla przedstawione są schematycznie na rys. 16.22. Kwadraty oznaczają dziury po jonach, kółka zaś — nowe ich położenia.



Rys. 16.22

Defekty sieci mogą też wynikać z istnienia w sieci domieszek. Tak np. KCl może mieć w niektórych węzłach sieci jon dwuwartościowy Ca^{++} zamiast jonu K^{+}, z tym jednak, że równocześnie powstaje luka po jonie Cl^{-}, co zapewnia zachowanie obojętnego stanu elektrycznego kryształu jako całości. Hipotezę utraty jonu Cl^{-} przy wprowadzeniu domieszki Ca^{++} potwierdzają bardzo dokładne pomiary gęstości kryształu.

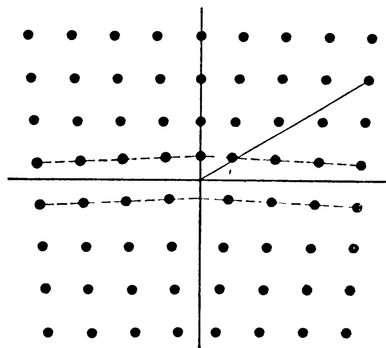
Defekty mogą być związane nie tylko z wprowadzeniem domieszek obcych, lecz również mogą wynikać z wprowadzenia np. nadmiaru kationu lub anionu w porównaniu z ilością wynikającą ze stosunków stechiometrycznych. Tak np. ogrzewanie kryształu NaCl w parze sodu powoduje wystąpienie domieszek Na, które sygnalizują swą obecność pojawieniem się żółtej barwy kryształu NaCl, przezroczystego — jak wiadomo — w świetle dziennym. Podobnie kryształ KCl ogrzewany w parze potasu uzyskuje mniej lub bardziej intensywną barwę czerwoną — zależnie od ciśnienia pary.

Pojawieniu się nadmiarowych jonów sodu w węzłach towarzyszy wytworzenie się luk w węzłach zajętych poprzednio przez jony Cl^{-}. Zniknięcie z sieci ładunku ujemnego $(-e)$ można traktować jako równoważne z pojawieniem się w tym wolnym miejscu ładunku dodatniego $(+e)$. W pobliżu luki zachowującej się tak, jakby miała ładunek dodatni, łatwo ulokowuje się elektron wyzwolony z nadmiarowego atomu Na przy jego

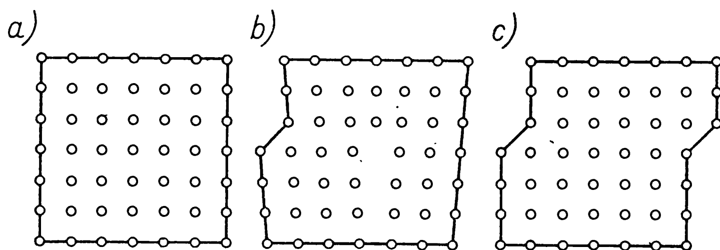
przekształceniu się w jon. Taki złożony defekt w postaci dodatniej dziury i pobliskiego elektronu ma zdolność pochłaniania światła w pewnym zakresie długości fali i dzięki tej absorpcji (por. § 28.6) powoduje wystąpienie zabarwienia kryształu. Ten związek obecności takiego defektu z pojawieniem się zabarwienia zdecydował o nazwie tego rodzaju defektu — *centrum F* (*F* — od niem. Farbe).

O wpływie domieszek będziemy jeszcze mówili w rozdz. 32, poświęconym fizyce ciała stałego. Tutaj tylko jeszcze podkreślimy, że defekty w krysztale nie mają stałych, określonych położeń. Tak np. do wytworzonej luki w sieci może przesunąć się sąsiedni jon lub atom wytwarzając lukę w miejscu swego poprzedniego położenia. Efekt ostateczny jest taki, jakby luka-defekt kryształu uległa przesunięciu.

Krótko wspomnimy jeszcze o defektach zwanych dyslokacjami, których ruch w krysztale ściśle się wiąże z właściwościami plastycznymi kryształu. Ograniczymy się do opisania prostej dyslokacji brzegowej.



Rys. 16.23



Rys. 16.24

Dyslokację brzegową ilustruje w dwóch wymiarach rys. 16.23. Jak widać, dyslokacja ta sprowadza się do wsunięcia do sieci półprostej (w trzech wymiarach — półpłaszczyzny) obsadzonej w węzłach elementami sieci. Wywołuje to zaburzenie w sieci: ściśnięcie atomów w jednej części kryształu, a rozsunięcie ich w drugiej. Zaburzenie wywołane ruchem takiej dyslokacji poprzez kryształ zilustrowane jest rys. 16.24. Po przejściu dyslokacji przez kryształ mamy efekt taki (rys. 16.24c), jakby odbył się poślizg górnej części kryształu względem dolnej o jedną odległość międzywęzłową.

Mówiąc o defektach kryształów można jeszcze wspomnieć o tzw. fononach. Będzie to tym bardziej wskazane, że przypomni nam o punkcie wyjściowym naszych rozważań — o założeniach teorii kinetyczno-molekularnej. Mówiąc o kryształach uwzględniliśmy ich budowę cząsteczkową, wymieniliśmy rodzaje wiązań krystalicznych, nie wspominaliśmy jednak o ruchu elementów budowy kryształów. Błędne byłoby wyobrażanie sobie atomów, jonów lub cząsteczek tkwiących nieruchomo w określonych punktach sieci. Elementy kryształu wykonują drgania. Energia tych drgań rośnie wraz z temperaturą. Przyrost tej energii odniesiony do jednostki masy i ogrzania o jeden stopień nosi nazwę ciepła właściwego. Ponieważ elementy kryształu są ze sobą powiązane pewnymi siłami, więc i ich drgania nie są od siebie niezależne. W wyniku drgań elementów powstają w krysztale fale o dużej częstotliwości, dające różne efekty falowe, jak interferencja, fale stojące itd. Te zjawiska można kojarzyć z przenoszeniem w krysztale porcji (paczek) energii zwanych *fononami*. Fononom przypisuje się analogicznie do fotonów energię równą iloczynowi stałej Plancka i częstotliwości oraz pęd $p = h\nu/v$, gdzie v oznacza prędkość fal sprężystych.

Na zakończenie jeszcze krótka uwaga o tzw. *ciekłych kryształach*. Nazwę tę nadaje się — zresztą niezbyt słusznie — cieczom, w których cząsteczki o długiej, pałeczkowatej strukturze wykazują w pewnych mikroskopijnych obszarach uporządkowanie sprowadzające się do równoległego ułożenia w przestrzeni podłużnych osi cząsteczek. To uporządkowanie decyduje o wystąpieniu pewnych cech charakterystycznych dla kryształów, np. podwójnego załamywania światła przy braku jednak innych, jak np. sprężystości postaciowej.

16.15. Wprowadzenie do teorii kinetycznej cieczy

Obserwując pod mikroskopem np. rozcieńczony roztwór tuszu, koloidalny roztwór złota lub zawiesinę cząstek gumiguty w wodzie można zauważyć przy dostatecznym powiększeniu bezładny ruch cząsteczek. Pierwszy raz taki ruch był obserwowany przez angielskiego botanika Browna w r. 1827. Ruchy tego typu nazywamy *ruchami Browna.*

Zazwyczaj ruchy Browna są omawiane jako wstęp do teorii kinetycznej cieczy, stanowią one bowiem bezpośredni, bardzo przekonywający dowód występowania ruchu cieplnego w cieczach. Ruchy Browna odbywają się niezależnie od tego, czy zawiesina jest świeżo zrobiona, czy też upłynął dłuższy czas od momentu jej sporządzenia. Usuwa to podejrzenie, że ruchy te są związane z brakiem ustalonej równowagi w cieczy, występującym bezpośrednio po przygotowaniu zawiesiny.

Obserwacje wykazują, że ze wzrostem temperatury wzrasta również prędkość ruchów Browna.

Obecnie rozważymy ruchy Browna jako potwierdzenie stosowalności zasady ekwipartycji energii do cząstek zarówno małych, jak i bardzo dużych. Omówimy też sposób wyznaczania wartości stałej Avogadra N_A na podstawie obserwacji ruchów Browna.

Ruchy Browna można traktować jako wynik zderzeń cząsteczek zawiesiny z cząsteczkami otaczającej cieczy. Dla uzyskania lepszego obrazu załóżmy początkowo, że za-

wiesinę stanowi ciało o dużych rozmiarach, zanurzone w cieczy. Ciało to w każdej chwili doznaje dużej liczby uderzeń ze strony cząsteczek otaczającej cieczy. Cząsteczki cieczy uderzają w powierzchnię ciała równomiernie ze wszystkich stron. Ze względu na dużą liczbę uderzeń w jednostce czasu i ich doskonałą symetrię w przestrzeni nie powodują one żadnej zmiany w położeniu ciała. W miarę zmniejszania się rozmiarów ciała maleje jego masa i powierzchnia. Zmniejsza się też liczba doznawanych uderzeń w jednostce czasu. W tych warunkach łatwiej może się zdarzyć, że zostanie zakłócona symetria uderzeń i że uderzenia cząsteczek cieczy z pewnego kierunku nie będą dokładnie kompensowane równoczesnymi uderzeniami cząsteczek z kierunku wprost przeciwnego. Wskutek tego ciało może zostać wprawione w ruch po torze zygzakowatym, tym częściej zmieniającym kierunek, im ciało jest lżejsze i mniejsze. Tak właśnie zachowują się cząsteczki zawiesin w cieczach. Dla orientacji podajemy, że cząstki zawiesin gumiguty mają średnice rzędu 1 mikrometra.

Einstein i Smoluchowski (niezależnie od siebie) zakładali, że zasada ekwipartycji energii stosuje się również do wielkich cząstek, jakimi są cząstki zawiesin. W oparciu o to założenie podali oni wzór, według którego można obliczyć średni kwadrat przesunięć $\overline{\Delta x^2}$ wykonanych przez daną cząstkę w czasie τ. Wzór ten ma postać

$$\overline{\Delta x^2} = \frac{RT\tau}{3N_A \eta \pi r}.$$

Jest on słuszny dla cząstek kulistych o promieniu r, znajdujących się w ośrodku o współczynniku lepkości η. R oznacza stałą gazową, T — temperaturę w skali Kelvina, N_A — stałą Avogadra.

Obserwując jakąś cząstkę zawiesiny pod mikroskopem można ustalić doświadczalnie wartość Δx_1 określającą wielkość przesunięcia cząstki obserwowanej w czasie τ (np. 30 sekund). Po następnych τ sekundach wielkość przesunięcia tej samej cząstki wynosi Δx_2 itd. Biorąc dostatecznie dużą liczbę zmierzonych przesunięć można obliczyć ich średni kwadrat $\overline{\Delta x^2}$. Powtarzając takie pomiary z różnymi zawiesinami w różnych ośrodkach i przy różnych temperaturach, a więc zmieniając parametry T, r i η stwierdzono doświadczalnie słuszność podanego wzoru. W ten sposób uzyskano potwierdzenie założenia Einsteina i Smoluchowskiego co do stosowalności zasady ekwipartycji do cząstek zawiesin.

Nawiązując raz jeszcze do wzoru Einsteina i Smoluchowskiego, stwierdzimy możliwość wyznaczenia stałej Avogadra. Pomiar stałej gazowej R nie nastręcza większych trudności. Średnicę cząstki łatwo też ustalić podczas obserwacji pod mikroskopem. W sposób wyżej opisany wyznaczamy $\overline{\Delta x^2}$. Jeśli znamy współczynnik lepkości cieczy η, mamy wszystkie dane do obliczenia wartości stałej Avogadra. Wyniki otrzymane w ten sposób są zgodne z wynikami osiągniętymi na innych drogach. Zgodność tych wyników potwierdza założenie, że średnie energie kinetyczne ruchu postępowego cząsteczek ośrodka i zawiesin są równe, mimo znacznych różnic ich mas i rozmiarów.

Ruchy Browna potraktowaliśmy jako skutek ruchu cieplnego cząsteczek cieczy. Podobnie zawiesiny występujące w gazach, np. cząstki dymu w powietrzu, pozwalają obserwować skutki ruchu cieplnego cząsteczek gazu (oczywiście ruch cieplny cząste-

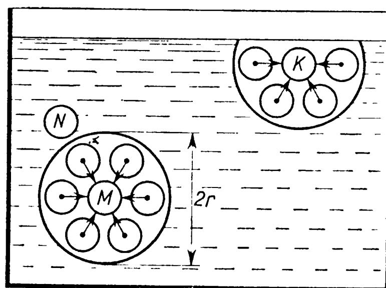
czek ośrodka istnieje i wtedy, gdy zawiesin nie ma — nie mamy jednak możności bezpośredniego stwierdzania istnienia tego ruchu).

Analogiczne rozważania, jak w przypadku zawiesin, odniesione do cząsteczek danego ośrodka, doprowadziły Smoluchowskiego do wniosku, że w cieczach i gazach muszą występować lokalne, chwilowe zmiany gęstości, tzw. *fluktuacje gęstości.* Zdaniem Smoluchowskiego fluktuacje gęstości są jedną z przyczyn rozpraszania światła w powietrzu, prowadzącego do powstania błękitu nieba. Odpowiednie doświadczenie wykonane przez Smoluchowskiego całkowicie potwierdziło jego hipotezę.

16.16. Ciśnienie molekularne w cieczach

Na podstawie teorii kinetyczno-molekularnej tłumaczymy szereg zjawisk zachodzących w pobliżu powierzchni swobodnej cieczy (na granicy faz).

Weźmy pod uwagę cząsteczkę M znajdującą się na pewnej głębokości pod powierzchnią swobodną cieczy (rys. 16.25). Niech r oznacza maksymalną odległość, przy której jeszcze występuje wzajemne przyciąganie cząsteczek. Nazwijmy tę odległość *promieniem sfery działania* cząsteczki. Promień sfery działania cząsteczki ma rząd wielkości 10^{-9} m. Wewnątrz kuli zakreślonej promieniem r ze środka cząsteczki M znajdują się wszystkie cząsteczki objęte działaniem cząsteczki M. Cząsteczka taka jak N już nie podlega działaniu cząsteczki M.



Rys. 16.25

Jeżeli cała sfera działania znajduje się pod powierzchnią swobodną cieczy, to wobec symetrii występowania cząsteczek w otoczeniu M wypadkowa siła działań międzycząsteczkowych na M równa się zeru.

Inaczej się sprawa przedstawia, jeśli część sfery działania znajduje się poza powierzchnią swobodną cieczy. Takie warunki zachodzą w odniesieniu do cząsteczki K na rys. 16.25. Symetria rozmieszczenia cząsteczek cieczy w sferze działania cząsteczki K jest zakłócona: wobec nadmiaru cząsteczek w półkuli dolnej cząsteczka K podlega sile wypadkowej skierowanej pionowo w dół. Siła ta usiłuje wciągnąć cząsteczkę K w głąb cieczy. Takiemu działaniu podlegają wszystkie cząsteczki znajdujące się na powierzchni swobodnej cieczy, jak również cząsteczki odsunięte od powierzchni swobodnej na odległość mniejszą od r. Siła wypadkowa odniesiona do cząsteczek cieczy związanych

z 1 m^2 jej powierzchni swobodnej wyraża liczbowo wartość tzw. ciśnienia molekularnego cieczy. Ciśnienie to jest skierowane w głąb cieczy, prostopadle do powierzchni swobodnej.

Wartości ciśnienia molekularnego w cieczach są olbrzymie, rzędu tysięcy atmosfer. Występowanie tak wielkich ciśnień molekularnych w cieczach tłumaczy ich małą ściśliwość. Zastosowanie ciśnienia zewnętrznego rzędu kilku czy nawet kilkudziesięciu atmosfer ma minimalny wpływ na objętość cieczy, której powierzchnia już jest poddana działaniu ciśnień molekularnych rzędu tysięcy atmosfer.

16.17. Napięcie powierzchniowe

Ustaliliśmy, że warstwa cieczy sięgająca od powierzchni swobodnej w głąb na odległość promienia sfery działania, czyli około 10 Å, posiada inne właściwości niż reszta cieczy. Omawiając zjawiska zachodzące na powierzchni swobodnej posługujemy się analogią między warstwą powierzchniową cieczy i napiętą sprężystą błoną. W przypadku równomiernego napięcia takiej błony na każdą jednostkę długości pomyślaną na powierzchni błony działa taka sama siła rozciągająca skierowana stycznie do powierzchni, a prostopadle do pomyślanego na powierzchni elementu długości. Działaniu tej siły zewnętrznej rozciągającej towarzyszy powstanie siły wewnętrznej sprężystej, dążącej do zmniejszenia powierzchni błony. Podobna siła sprężysta **F**, styczna do powierzchni swobodnej cieczy, działa na każdy element długości L pomyślany na tej powierzchni. Stosunek siły F do długości elementu L nosi nazwę *współczynnika napięcia powierzchniowego* σ:

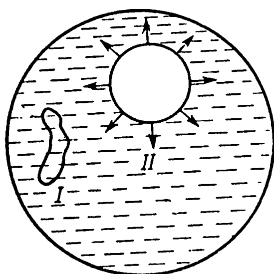
$$\sigma = \frac{F}{L}, \quad \text{czyli} \quad F = \sigma L. \tag{16.17}$$

Siłę F często nazywa się *siłą napięcia powierzchniowego.* Jest ona styczna do powierzchni swobodnej cieczy, a prostopadła do rozważanego elementu długości L.

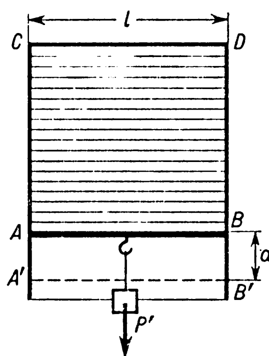
Należy podkreślić, że analogii między napiętą sprężystą błoną i powierzchnią swobodną cieczy nie należy traktować dosłownie. Jeśli napięta sprężysta błona ulega skurczeniu, to jej napięcie maleje. Jeśli natomiast zmniejsza się powierzchnia swobodna cieczy, to napięcie powierzchniowe nie ulega zmianie, gdyż współczynnik napięcia powierzchniowego jest wielkością charakterystyczną dla danej cieczy w danej temperaturze, a nie zależy od wielkości powierzchni swobodnej.

Występowanie sił napięcia powierzchniowego tłumaczy ogólnie znaną dążność cieczy do zmniejszania powierzchni swobodnej. Ten skutek istnienia napięcia powierzchniowego na powierzchni płaskiej można wykazać posługując się cienkim pierścieniem z drutu. Po wyciągnięciu pierścienia z mydlanej wody pozostaje na nim błonka mydlana, czyli cienka płaska warstwa cieczy ograniczona z obu stron powierzchniami swobodnymi. Rzućmy na tę błonkę nitkę związaną w pętlę (rys. 16.26). Nitka utrzymuje się na powierzchni błonki zachowując dowolnie wygięty kształt (I). Wystarczy jednak przekłuć błonkę wewnątrz pętli, aby spowodować rozwinięcie się pętli w koło (II). Siły napięcia powierzchniowego usiłują zmniejszyć powierzchnię swobodną błonki i tak ściągają nitkę stanowiącą brzeg błonki, że przyjmuje ona kształt koła.

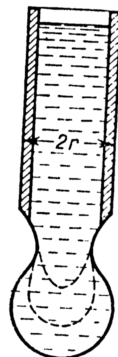
Wartość współczynnika napięcia powierzchniowego można wyznaczyć w następującym doświadczeniu. Prostokątna ramka metalowa *ABCD* (rys. 16.27) ma bok *AB* możliwie lekki i przesuwający się prawie bez tarcia wzdłuż *CA* i *DB*. Przez zanurzenie ramki do mydlanej wody otrzymujemy na ramce napiętą mydlaną błonkę. Pod działaniem siły napięcia powierzchniowego ruchomy bok ramki przesuwa się w stronę *CD*.



Rys. 16,26



Rys. 16.27



Rys. 16.28

W ten sposób zmniejsza się powierzchnia swobodna błonki. Obciążając bok *AB* odpowiednim odważnikiem można zapobiec jego przesuwaniu się. W tych warunkach siła *P*, równa ciężarowi odważnika i poprzeczki *AB*, równoważy całkowitą siłę napięcia powierzchniowego. Wartość tej siły znajdziemy mnożąc współczynnik napięcia powierzchniowego σ przez $2l$, gdzie l oznacza długość boku *AB* ramki. Czynnik 2 wiąże się z tym, że błonka napięta na ramce ma dwie powierzchnie swobodne. A zatem

$$P = 2l\sigma,$$

skąd

$$\sigma = \frac{P}{2l}. \tag{16.18}$$

Inny sposób pomiaru współczynnika napięcia powierzchniowego polega na badaniu wypływu cieczy oddzielnymi kroplami (rys. 16.28). Tuż przed oderwaniem się kropli siła napięcia powierzchniowego równoważy się z ciężarem *P* kropli. Siła napięcia powierzchniowego działa na obwodzie, wzdłuż którego kropla przylega do wylotu rurki o promieniu zewnętrznym *r*. Spełnione jest równanie

$$P = 2\pi r\sigma.$$

Ciężar *P* kropli znajdujemy odliczając *n* kropel i dzieląc ich łączny ciężar przez *n*. Dokładniejszy wzór otrzymujemy wprowadzając współczynnik poprawkowy, którego wartość liczbowa zależna jest od stosunku $r/V^{1/3}$, gdzie *V* oznacza objętość kropli. Wprowadzamy ten współczynnik ze względu na to, że w chwili odrywania się kropli pojawia się przewężenie i długość obwodu przylegania należy w odpowiednim stosunku zmniejszyć. Przykładowe wartości współczynników napięcia powierzchniowego podaje tab. 16.5.

Tabela 16.5

Współczynniki napięcia powierzchniowego

Rodzaj cieczy	Temperatura °C	σ 10^{-3} J/m²
eter etylowy	20	17
alkohol etylowy	20	23
benzen	20	29
woda	0	75,6
	10	74,2
	20	72,7
	100	58,9
rtęć	20	480,0
złoto	1130	1100,0

16.18. Energia potencjalna warstwy powierzchniowej cieczy

Każda z cząstek warstwy powierzchniowej podlega działaniu siły skierowanej w głąb cieczy. Stąd wniosek, że przesunięcie cząstki z głębi cieczy do warstwy powierzchniowej wymaga wkładu pracy. Praca włożona pozostaje w cieczy pod postacią przyrostu energii potencjalnej warstwy powierzchniowej.

Powróćmy do doświadczenia z ramką o ruchomym boku (rys. 16.27). Niech działanie minimalnej nadwyżki ciężaru P nad siłami napięcia powierzchniowego wywoła przesunięcie boku AB o odległość d. Towarzyszy temu wzrost powierzchni swobodnej błonki. Praca W siły zewnętrznej P określa przyrost energii potencjalnej błonki ΔE_p związany ze zmianą jej powierzchni swobodnej

$$W = \Delta E_p.$$

Praca W zużywa się na pokonanie oporu sił napięcia powierzchniowego, czyli

$$\Delta E_p = 2l\sigma d.$$

Iloczyn ld oznacza zmianę pola każdej z obu powierzchni swobodnych błonki. Iloczyn $2ld$ określa zatem całkowitą zmianę ΔS powierzchni swobodnej

$$\Delta E_p = \sigma \Delta S,$$

skąd

$$\sigma = \frac{\Delta E_p}{\Delta S}. \tag{16.19}$$

Na podstawie tego równania możemy podać drugą definicję współczynnika napięcia powierzchniowego. *Współczynnik napięcia powierzchniowego* odpowiada liczbowo wzrostowi energii potencjalnej warstwy powierzchniowej cieczy związanemu z powiększeniem powierzchni swobodnej o jednostkę. Z równania (16.19) wynika, że jednostką współczynnika napięcia powierzchniowego w układzie SI jest J/m². Łatwo się przekonać, że ta jednostka jest zgodna z jednostką poprzednio ustaloną — N/m.

Na tle tego określenia współczynnika napięcia powierzchniowego widać, że występująca w cieczach dążność do zmniejszania powierzchni swobodnej związana jest z przechodzeniem cieczy do takiego stanu, któremu odpowiada mniejsza wartość energii potencjalnej. W tabeli 16.5 podane są wartości współczynników napięcia powierzchniowego różnych cieczy. Wartości te zależą nie tylko od rodzaju cieczy, lecz także od temperatury. Ze wzrostem temperatury współczynnik napięcia powierzchniowego maleje i staje się równy zeru w temperaturze krytycznej, w której znika różnica między cieczą i jej parą nasyconą.

Podkreślamy raz jeszcze, że współczynnik napięcia powierzchniowego nie zależy od wielkości powierzchni swobodnej cieczy, jest on bowiem miarą energii potencjalnej warstwy powierzchniowej cieczy, przypadającej na jednostkę powierzchni swobodnej.

Na powierzchni cieczy można obserwować występowanie paradoksalnych zjawisk. Tak np. igła położona ostrożnie, płasko na powierzchni wody nie tonie, lecz pływa, mimo że jej ciężar właściwy jest większy od ciężaru właściwego wody. Tonięciu zapobiega błonka powierzchniowa związana z napięciem powierzchniowym.

Podobnie nie tonie w wodzie kawałek kamfory mającej również ciężar właściwy większy od ciężaru właściwego wody. Kamfora umieszczona na powierzchni wody wykonuje nieregularne ruchy (tzw. taniec kamfory) związane z tworzeniem się nietrwałego tzw. roztworu powierzchniowego: coraz to nowe cząsteczki kamfory wydostają się na powierzchnię i ulatniają. Podobne roztwory powierzchniowe, ale już trwałe, dają kwasy tłuszczowe, np. kwas oleinowy, stearynowy itd. Dość szeroko znany jest roztwór powierzchniowy w postaci rozprzestrzeniającej się na powierzchni wody cienkiej warstwy oliwy.

Wspomniane zjawiska dawniej tłumaczono różnicą współczynników napięcia powierzchniowego stykających się cieczy (woda-roztwór kamfory, woda-oliwa).

Krople cieczy poddane tylko działaniu napięcia powierzchniowego przyjmują dokładnie kształt kulisty, kula bowiem jest taką bryłą, która przy danej objętości ma najmniejszą powierzchnię. Obserwowane zwykle przez nas krople cieczy nie mają dokładnie kształtu kuli: wiąże się to z zakłócającym działaniem dodatkowych sił, najczęściej siły ciężkości. Wpuśćmy jednak kroplę oleju do mieszaniny wody i alkoholu tak sporządzonej, by ciężar właściwy mieszaniny równał się ciężarowi właściwemu oleju. W tych warunkach siła ciężkości jest całkowicie zrównoważona przez siłę wyporu. Kropla oleju pod działaniem sił napięcia powierzchniowego przyjmuje dokładnie kształt kuli.

16.19. Zależność ciśnienia molekularnego od kształtu powierzchni swobodnej

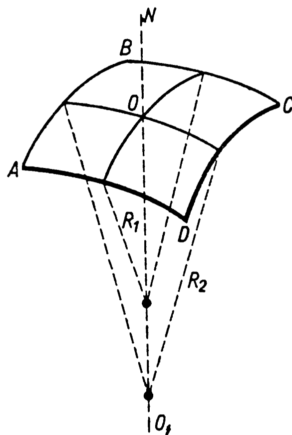
Mówiąc o właściwościach warstwy powierzchniowej cieczy wprowadziliśmy pojęcie ciśnienia molekularnego cieczy. Nasuwa się pytanie, czy wartość tego ciśnienia jest zależna od kształtu powierzchni swobodnej?

Na początku ubiegłego stulecia Laplace podał teorię i wzory określające wartość ciśnienia molekularnego w zależności od krzywizny powierzchni. Oznaczmy ciśnienie molekularne na płaskiej powierzchni swobodnej przez p_0. Według Laplace'a ciśnienie

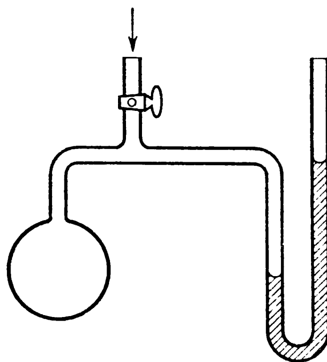
molekularne p na zakrzywionej (wklęsłej lub wypukłej) powierzchni cieczy można wyrazić wzorem

$$p = p_0 + \sigma\left(\frac{1}{R_1} + \frac{1}{R_2}\right),$$

gdzie σ oznacza współczynnik napięcia powierzchniowego, R_1 i R_2 — główne promienie krzywizny danej powierzchni. *Główne promienie krzywizny* możemy określić w ten sposób: niech $ABCD$ (rys. 16.29) przedstawia część powierzchni zakrzywionej. W punkcie O tej powierzchni wystawiamy normalną ON. Przez tę normalną można poprowadzić nieskończenie wiele par płaszczyzn przecięcia wzajemnie do siebie prostopadłych. Każdej z płaszczyzn przecięcia odpowiada na powierzchni badanej pewna krzywa o określonym promieniu krzywizny. Można tak dobrać parę płaszczyzn przecięcia, aby wspomniane promienie krzywizny miały wartości ekstremalne: jeden — wartość największą, drugi — wartość najmniejszą z możliwych. Takie promienie krzywizny nazywamy głównymi promieniami krzywizny danej powierzchni.



Rys. 16.29



Rys. 16.30

Łatwo sprawdzić, że dla powierzchni płaskiej ($R_1 = R_2 = \infty$) wzór Laplace'a redukuje się do wyrażenia

$$p = p_0.$$

Na powierzchni wypukłej oba promienie krzywizny są dodatnie, czyli

$$p > p_0.$$

Ciśnienie molekularne na powierzchni wypukłej jest większe niż na powierzchni płaskiej. Innymi słowy, od strony wypukłej powierzchni swobodnej działa na ciecz dodatkowe ciśnienie wewnętrzne, równe

$$\sigma\left(\frac{1}{R_1} + \frac{1}{R_2}\right).$$

Na powierzchni wklęsłej mamy sytuację przeciwną, gdyż oba promienie krzywizn są ujemne, a więc

$$p < p_0,$$

czyli od wklęsłej strony powierzchni swobodnej ciecz jest poddana ciśnieniu skierowanemu od cieczy na zewnątrz. W wypadku szczególnym, gdy badana powierzchnia jest kulista, $R_1 = R_2 = R$,

$$p = p_0 + \frac{2\sigma}{R}.$$

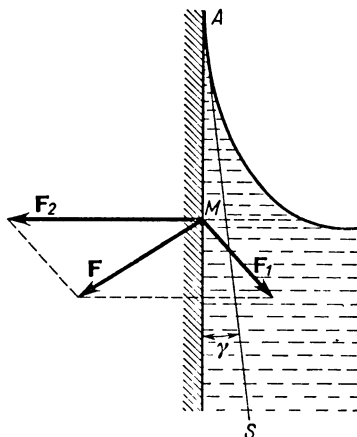
Jeśli np. rozważamy ciśnienie panujące we wnętrzu kulistej bańki mydlanej, to musimy uwzględnić dwie powierzchnie swobodne bańki: zewnętrzną — wypukłą i wewnętrzną — wklęsłą. Od powierzchni swobodnej zewnętrznej działa na ciecz ciśnienie dodatkowe, skierowane w stronę cieczy, czyli do środka bańki. Na wklęsłej powierzchni swobodnej działa na ciecz ciśnienie skierowane od cieczy na zewnątrz. Ale ten kierunek „od cieczy na zewnątrz" odniesiony do wewnętrznej powierzchni bańki mydlanej jest kierunkiem do wnętrza bańki. Innymi słowy, na powietrze zawarte w bańce mydlanej działa dodatkowe ciśnienie, wynikające z zakrzywienia powierzchni cieczy, i równe $4\sigma/R$ (ciśnienia p_0 na obu powierzchniach swobodnych bańki są równe co do wartości, lecz przeciwnie skierowane, a zatem równoważą się wzajemnie w wypadkowym działaniu). To ciśnienie równoważy nadwyżkę ciśnienia istniejącą w bańce, związaną z jej wydmuchaniem. Tę nadwyżkę ciśnienia w bańce można łatwo stwierdzić wydmuchując bańkę na rurce szklanej dołączonej do manometru wodnego (rys. 16.30).

Jeśli zamiast bańki mydlanej, zawierającej w środku powietrze, weźmiemy pod uwagę kulistą kroplę cieczy, będziemy mieli do czynienia z jedną tylko powierzchnią swobodną — wypukłą. W związku z krzywizną wystąpi na tej powierzchni dodatkowe ciśnienie wewnętrzne skierowane do wnętrza kropli i równe $2\sigma/R$.

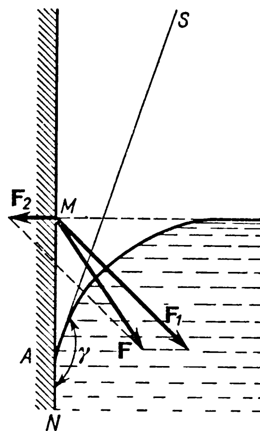
16.20. Przyleganie

Interesujące zjawiska powierzchniowe występują w miejscach zetknięcia ciał stałych i cieczy. Z obserwacji codziennych wiemy, że powierzchnie swobodne cieczy spoczywających w naczyniach szerokich można traktować jako powierzchnie poziome i płaskie. Jednak w najbliższym otoczeniu ścianek naczynia występuje zakrzywienie powierzchni najwyraźniej dające się zaobserwować w rurkach o małym przekroju, rzędu 1 mm^2, zwanych rurkami kapilarnymi albo włoskowatymi. Powierzchnia swobodna cieczy w takich rurkach jest wypukła lub wklęsła, zależnie od rodzaju cieczy i materiału, z jakiego sporządzona jest rurka. Mówimy w tych przypadkach o *menisku wypukłym lub wklęsłym*. Tak np. menisk wody lub nafty w naczyniu szklanym jest wklęsły. Menisk rtęci w rurce szklanej jest wypukły, menisk rtęci w rurce miedzianej, srebrnej lub złotej jest wklęsły itd. Jak można wyjaśnić różne zachowanie się tych cieczy w pobliżu ścian naczynia?

Rozważmy cząsteczkę M cieczy, znajdującą się na powierzchni swobodnej przy ściance naczynia (rys. 16.31 i 16.32). Załóżmy, początkowo, że menisk cieczy jest płaski. Wtedy wybrana cząsteczka podlega działaniu sił przyciągających ze strony cząsteczek tej samej cieczy znajdujących się w $\frac{1}{4}$ sfery działania. Wypadkowa tych sił $\mathbf{F}_1$ skierowana jest w głąb cieczy pod kątem 45° względem powierzchni swobodnej cieczy. Równocześnie cząsteczka M jest przyciągana siłą $\mathbf{F}_2$ przez cząsteczki ciała stałego (ścianki naczynia) w kierunku prostopadłym do ścianki. Ponadto występuje siła oddziaływania cząsteczek gazu znajdującego się nad cieczą oraz siła ciężkości cząsteczki M, ale obie te siły można zaniedbać jako bardzo małe w porównaniu z pozostałymi siłami.



Rys. 16.31



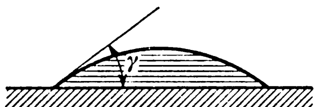
Rys. 16.32

Siły $\mathbf{F}_1$ i $\mathbf{F}_2$ dają wypadkową $\mathbf{F}$, której kierunek zależny jest od względnych wartości sił $\mathbf{F}_1$ i $\mathbf{F}_2$. Jeżeli siła „przylegania" do ciała stałego jest stosunkowo duża w porównaniu z siłą „spójności" między cząsteczkami cieczy, to wypadkowa jest skierowana tak, jak na rys. 16.31. Przypadek przeciwny ilustruje rys. 16.32. Wiemy jednak, że ciecz pozostaje w spoczynku jedynie wtedy, gdy siły wypadkowe, działające na cząsteczki cieczy na powierzchni swobodnej, są prostopadłe do tej powierzchni. Innymi słowy, cząsteczki cieczy w otoczeniu M przesuwają się tak, aby utworzyla się powierzchnia swobodna prostopadła do kierunku siły wypadkowej działającej na poszczególne cząsteczki. Odpowiada to wznoszeniu się cieczy wzdłuż ścianki (rys. 16.31). Powstaje przy tym menisk wklęsły. Ciecz, która tak się zachowuje, nazywamy *cieczą zwilżającą* dane ciało stałe.

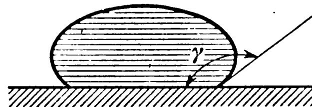
Na rysunku 16.32 widzimy menisk wypukły. Znów powierzchnia swobodna cieczy po ustaleniu się równowagi jest prostopadła do kierunku siły wypadkowej. Mówimy, że ciecz tworząca menisk wypukły jest *cieczą niezwilżającą* danego ciała stałego. Uzupełnijmy rys. 16.31 i 16.32 poprowadzeniem z punktu A stycznej do zakrzywionej powierzchni cieczy. Kąt MAS (rys. 16.31), jak również kąt NAS (rys. 16.32) nazywamy *kątem zetknięcia* i oznaczamy literą γ. Jest to kąt ostry w przypadku cieczy zwilżającej,

a rozwarty w przypadku cieczy niezwilżającej. Wielkość tego kąta jest stała w danej temperaturze dla określonego układu ciecz-ciało stałe. Zmiana nachylenia ścianek naczynia względem pionu nie ma wpływu na wielkość tego kąta. Jeśli kroplę cieczy umieścimy na poziomej płytce, to kształt, jaki kropla przyjmie, będzie zależny od kąta zetknięcia charakterystycznego dla danej cieczy i ciała stałego. Na rys. 16.33 widzimy kroplę cieczy zwilżającej, a na rys. 16.34 — kroplę cieczy niezwilżającej materiału płytki.

Takie ciecze, które rozpływają się po powierzchni ciała stałego nie osiągając położenia równowagi (np. pełzną coraz wyżej po ściankach naczynia), nazywamy *doskonale zwilżającymi*. Dla takich cieczy kąt zetknięcia z danym ciałem stałym równa się zeru.



Rys. 16.33



Rys. 16.34

Zjawisko zwilżania, ewentualnie niezwilżania ciała stałego przez daną ciecz, wykorzystuje się na wielką skalę w procesie oddzielania rud metalicznych od jałowego podłoża. Mieszaninę zmielonej rudy i podłoża wprowadza się w szybki ruch w takiej cieczy, która jest zwilżająca dla rudy, lecz niezwilżająca dla podłoża. Powietrze przepuszczane przez taką ciecz osiada w postaci pęcherzyków na niezwilżonych cząstkach podłoża i powoduje ich wypływanie na powierzchnię. Zwilżone cząstki rudy opadają na dno zbiornika. Opisany proces nosi nazwę *flotacji*.

Warto podkreślić, że omówione wyżej siły przylegania występują nie tylko na granicy zetknięcia ciecz-ciało stałe. Tym samym mianem obejmujemy siły międzycząsteczkowe, występujące między cząsteczkami ciała stałego lub gazu i cząsteczkami innego ciała stałego. Tak więc utrzymywanie się kredy na tablicy podczas pisania będzie też wynikiem sił przylegania.

16.21. Włoskowatość

Można zrobić interesujące spostrzeżenia zanurzając rurkę włoskowatą w szerszym naczyniu z cieczą. Otrzymujemy naczynie połączone o dwóch ramionach, w których — w przypadku równowagi — ciecz powinna się znajdować na wspólnym poziomie. Tymczasem doświadczenie wykazuje, że ciecz zwilżająca w rurce włoskowatej ma poziom wyższy od poziomu cieczy w szerszym naczyniu, a ciecz niezwilżająca — poziom niższy. To podniesienie poziomu cieczy zwilżającej, jak również obniżenie poziomu cieczy niezwilżającej, występujące szczególnie wyraźnie w rurkach włoskowatych, obejmujemy mianem *włoskowatości*.

Przyjmijmy, że powierzchnia swobodna cieczy w rurce włoskowatej jest częścią powierzchni kulistej o promieniu R. Zgodnie z tym, co mówiliśmy w związku z prawem Laplace'a, na menisku wklęsłym działa na ciecz ciśnienie skierowane od cieczy na zewnątrz (a więc do góry), równe $2\sigma/R$. Dla zrównoważenia tego ciśnienia poziom cie-

czy podnosi się o h cm. Ciśnienie hydrostatyczne odpowiadające wysokości h równa się wartości ciśnienia $2\sigma/R$. Stąd

$$\frac{2\sigma}{R} = h\varrho g.$$

Na rys. 16.35 widać, że

$$R\cos\gamma = r,$$

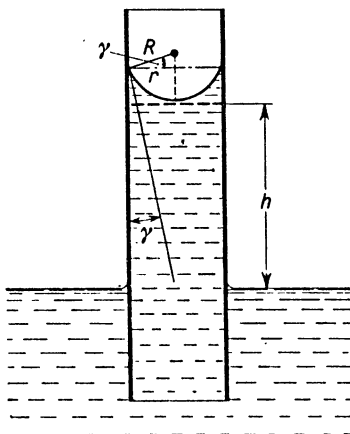
gdzie r jest promieniem rurki włoskowatej. A zatem

$$\frac{2\sigma\cos\gamma}{r} = h\varrho g,$$

skąd

$$h = \frac{2\sigma\cos\gamma}{r\varrho g}.$$

W ten sposób można obliczyć obniżenie włoskowate w cieczy niezwilżającej. Jak widać ze wzoru, przesunięcia włoskowate różnych cieczy są proporcjonalne do współczynnika napięcia powierzchniowego oraz do cosinusa kąta zetknięcia i odwrotnie proporcjonalne do promienia rurki włoskowatej oraz gęstości cieczy.



Rys. 16.35

Zjawiska włoskowatości odgrywają bardzo ważną rolę zarówno w przyrodzie, jak i w życiu praktycznym. Wystarczy wymienić wsiąkanie wody w ziemię lub jakiekolwiek inne porowate ciało, podnoszenie się wody do góry w cienkich włóknach roślinnych wbrew działaniu siły ciężkości, rolę włoskowatości w krążeniu krwi w organizmach żywych itd.

Pytania i zadania

1. Jakie są podstawowe założenia teorii kinetyczno-molekularnej?
2. Dlaczego rozwinięcie teorii kinetycznej ciał stałych, ciekłych i gazowych napotyka różne trudności?
3. Jaki jest przebieg zmian siły i energii potencjalnej, związanych z oddziaływaniami międzycząsteczkowymi, w zależności od odległości międzycząsteczkowych?

4. Wymienić, zdefiniować i omówić rzędy wartości liczbowych podstawowych wielkości z teorii kinetycznej gazów.

5. Jaka istnieje zależność między prędkością średnią, średnią kwadratową i najprawdopodobniejszą?

6. Jaki jest przebieg krzywej rozkładu prędkości Maxwella? Jakie wielkości odłożone są na osiach współrzędnych? Jaki jest ich sens fizyczny? Jakie wnioski można wyciągnąć z przebiegu krzywej?

7. Podać szczegółowy plan wyprowadzenia wzoru na ciśnienie gazu według teorii kinetycznej gazów Jakie są ostateczne postacie wzoru?

8. W jaki sposób w oparciu o teorię kinetyczną gazów można uzasadnić prawo Boyle'a–Mariotte'a. Avogadra i Daltona?

9. Jaka jest kinetyczna definicja temperatury?

10. Omówić sens fizyczny pojęcia liczba stopni swobody. Czy liczba stopni swobody cząsteczek wieloatomowych we wszystkich temperaturach jest ściśle określona? Czy stopnie swobody wiążą się tylko z ruchami postępowymi i obrotowymi cząsteczek?

11. Podać treść zasady ekwipartycji energii. Czy średnia energia kinetyczna, przypadająca na jedną cząsteczkę helu w temperaturze bliskiej zera bezwzględnego i w temperaturze pokojowej, jest jednakowa?

12. W jaki sposób od energii wewnętrznej molowej można przejść do pojemności cieplnej jednego mola przy stałej objętości (C_V)?

13. Jakie wnioski można wyciągnąć z faktu, że ciepło molowe wodoru w temperaturze około 100K wynosi około 12,5 J/mol · K, a w temperaturze około 400K — około 20,9 J/mol · K?

14. Jakie wspólne cechy posiadają trzy poznane zjawiska transportu? Jakie wielkości fizyczne ulegają transportowi w zjawisku przewodnictwa cieplnego, dyfuzji i lepkości? Dzięki czemu odbywa się transport tych wielkości?

15. Które z poznanych współczynników określających zjawiska transportu zależą od średniej drogi swobodnej cząsteczek, ciśnienia i temperatury?

16. Jakie są podstawowe formy występowania ciał stałych? Jaka jest różnica między ciałem krystalicznym i ciałem bezpostaciowym?

17. Jaki podział kryształów można wprowadzić ze względu na: a) geometryczną prawidłowość budowy, b) siły wiązań międzycząsteczkowych?

18. Scharakteryzować różne typy sił wiążących w krysztale.

19. Jakie są najczęściej spotykane defekty krystaliczne? Podać ich charakterystyki. Do czego się sprowadza dyslokacja brzegowa?

20. Jakie zjawisko nosi nazwę ruchów Browna? Jak można je tłumaczyć?

21. Co nazywamy sferą działania cząsteczki?

22. Co powoduje występowanie w cieczach ciśnienia molekularnego?

23. Scharakteryzować zjawisko napięcia powierzchniowego. Podać dwie definicje współczynnika napięcia powierzchniowego. Omówić jednostki współczynnika napięcia powierzchniowego w układzie SI. Jaki jest rząd wielkości tych współczynników?

24. Czym się tłumaczą różne kształty powierzchni swobodnych cieczy w różnych naczyniach? Ile wynosi kąt zwilżania dla cieczy doskonale zwilżającej?

25. Wyjaśnić zjawisko włoskowatości. Podać jego różne przykłady. Wyjaśnić dzięki czemu pomiar wzniesienia lub obniżenia włoskowatego można zastosować do wyznaczania współczynnika napięcia powierzchniowego?

26. W jakiej temperaturze prędkość średnia kwadratowa cząsteczek dwutlenku węgla będzie równa prędkości średniej kwadratowej cząsteczek azotu w temperaturze 0°C?

Odp. 156°C.

27. Ile cząsteczek znajduje się w 1 l gazu pod ciśnieniem 0,01 mm Hg w temperaturze $t = 7°C$?

Odp. ok. $3 \cdot 10^{17}$.

28. Jaka temperatura odpowiada prędkości średniej kwadratowej cząsteczek dwutlenku węgla równej 720 km/h?

Odp. −203°C.

29. Mieszanina gazów składająca się z 12 g tlenu i 24 g dwutlenku węgla znajduje się w zbiorniku o pojemności 3 l w temperaturze 127°C. Obliczyć ciśnienie mieszaniny gazów.

Odp. ok. 5 atm.

30. Obliczyć średnią energię kinetyczną ruchu postępowego atomów we wnętrzu Słońca, jeżeli wiadomo, że temperaturę wnętrza szacuje się na ok. $2 \cdot 10^7$K.

Odp. ok. $4 \cdot 10^{-16}$ J.

31. Przy jakim ciśnieniu w temperaturze 0°C średnia droga swobodna cząsteczek azotu będzie równa 1 mm, jeżeli w warunkach normalnych wynosi ona $6 \cdot 10^{-6}$ cm?

Odp. $4{,}6 \cdot 10^{-2}$ mm Hg.

32. Wiedząc, że średnia droga swobodna cząsteczek H_2 w warunkach normalnych wynosi $l_0 = 1{,}28 \cdot 10^{-5}$ cm obliczyć: a) drogę swobodną w temperaturze 0°C i pod ciśnieniem 10^{-4} mm Hg oraz b) średnicę cząsteczki.

33. Obliczyć średnią prędkość cząsteczek argonu, jeśli wiadomo, że średni czas między dwoma kolejnymi zderzeniami cząsteczek wynosi $9 \cdot 10^{-7}$ s i liczba cząsteczek w 1 cm³ wynosi $3{,}4 \cdot 10^{12}$. Średnia droga swobodna cząsteczek argonu w warunkach normalnych wynosi $6{,}66 \cdot 10^{-6}$ cm.

Odp. $586 \cdot 10^3$ m/s.

34. Jaką pracę należy wykonać, aby kroplę oleju o masie 1 g rozbić w wodzie na kropelki o średnicach równych $2 \cdot 10^{-4}$ cm, jeżeli: a) proces zachodzi izotermicznie, b) napięcie powierzchniowe na granicy woda-olej równa się $18 \cdot 10^{-3}$ N/m, c) gęstość oleju wynosi 0,9 g/cm³.

Odp. 0,06 J.

35. Rurka włoskowata o średnicy wewnętrznej $d = 1{,}5$ mm i długości $l = 20$ cm została w ten sposób zanurzona w położeniu poziomym w rtęci na głębokości $h = 10$ cm, że w rurce pozostała cała masa zawartego w niej pierwotnie powietrza. Jaka będzie długość słupka powietrza w rurce, jeżeli ciśnienie zewnętrzne będzie ciśnieniem normalnym?

Odp. 17,8 cm.

ROZDZIAŁ 17

Przemiany fazowe

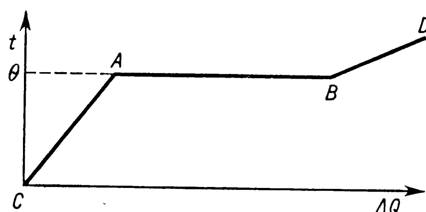
Głównym tematem niniejszego rozdziału są procesy, zwane *przemianami fazowymi*, tzn. takie procesy, w których w sposób gwałtowny, skokowy ulegają zmianie różne właściwości ciała. Do przemian fazowych zaliczamy topnienie i krzepnięcie (ciał krystalicznych), parowanie i skraplanie, sublimację itp. Z tych przykładów widać, że pojęcie przejścia fazowego można kojarzyć ze zmianą stanu skupienia: stan stały, ciekły i gazowy — są to różne fazy danej substancji. *Pojęcie fazy jest jednak pojęciem szerszym niż pojęcie stanu skupienia*: określona substancja w jednym stanie skupienia może występować — jak zobaczymy — w różnych fazach.

17.1. Zjawisko topnienia i krzepnięcia ciał

Wiemy już, że jeżeli skutkiem doprowadzania ciepła do badanego ciała jest wzrost jego temperatury, to obowiązuje zależność

$$\Delta Q = m\bar{c}\Delta T, \tag{17.1}$$

gdzie m oznacza masę ciała, $\bar{c}$ — ciepło właściwe średnie, ΔT — przyrost temperatury. Ze wzrostem temperatury na ogół stopniowo rosną wymiary ciała, co wiąże się ze zwięk-



Rys. 17.1

szoną amplitudą drgań atomów (cząsteczek lub jonów w sieci krystalicznej), stopniowo ulegają zmianie moduły sprężystości itp. Jeżeli wyniki równomiernego ogrzewania danego ciała przedstawimy na wykresie, odkładając na osi odciętych ilość doprowadzanego ciepła (czas równomiernego ogrzewania), a na osi rzędnych temperaturę, to otrzymamy odcinek linii prostej (AC na rys. 17.1).

Kontynuując ogrzewanie badanego ciała krystalicznego dochodzimy wreszcie do takiej temperatury, charakterystycznej dla danego ciała przy danym ciśnieniu zewnętrznym, przy której rozpoczyna się *proces topnienia*. Temperaturę tę nazywamy *temperaturą topnienia*. Od chwili osiągnięcia temperatury topnienia — mimo ciągłego doprowadzania ciepła — temperatura ciała nie ulega dalszej zmianie. Energia doprowadzana z zewnątrz w postaci ciepła wywołuje gwałtowną zmianę struktury wewnętrznej ciała, połączoną z zanikiem sieci krystalicznej. Gdy całe ciało ulegnie stopieniu, następuje dalszy wzrost temperatury, ale on już dotyczy ciała w fazie ciekłej. Na rys. 17.1 odcinek AB przedstawia przebieg procesu topnienia w stałej temperaturze topnienia θ. Odcinek BD odpowiada ogrzewaniu cieczy. Bardziej stromy przebieg odcinka AC niż odcinka BD przypomina, że dane ciało w fazie stałej ma na ogół mniejsze ciepło właściwe niż powstająca z niego ciecz.

Podczas odbierania ciepła od ciała ciekłego otrzymalibyśmy te same punkty wykresu, tylko w przeciwnej kolejności: od D przez B, A do C. Tym razem odcinek BA odpowiadałby *procesowi krzepnięcia*.

Jeżeli po rozpoczęciu procesu topnienia przerwiemy dopływ ciepła, stwierdzimy, że w otrzymanej mieszaninie dwóch faz nic się nie dzieje: nie zachodzi ani proces topnienia, ani proces krzepnięcia. Osiągnięta *temperatura topnienia* jest — przy danym ciśnieniu zewnętrznym — temperaturą równowagi fazy ciekłej i fazy stałej.

Utrzymywanie procesu topnienia wymaga doprowadzania ciepła do ciała ogrzanego do temperatury topnienia. Ilość ciepła ΔQ potrzebna do stopienia masy m danego ciała stałego w temperaturze topnienia jest proporcjonalna do tej masy:

$$\Delta Q = mc_t, \tag{17.2}$$

gdzie c_t oznacza współczynnik proporcjonalności, zwany ciepłem topnienia danego ciała.

Ciepło topnienia danego ciała wyraża liczbowo ilość energii cieplnej, potrzebną do stopienia 1 kg danego ciała w temperaturze topnienia. Ciepło topnienia w układzie SI wyraża się w J/kg.

Metoda wyznaczania ciepła topnienia, np. za pomocą kalorymetru wodnego, opiera się na zasadzie bilansu cieplnego. Przypomni tę metodę przeanalizowanie podanego niżej przykładu.

P r z y k ł a d. Do mosiężnego kalorymetru wodnego o masie m_k i cieple właściwym c_k, zawierającego masę m_w wody o cieple właściwym c_w i temperaturze t_0, wrzucono masę m_x badanego ciała krystalicznego o cieple właściwym c_x' i temperaturze t. Temperatura topnienia θ ciała jest mniejsza od temperatury t_0 i od temperatury końcowej t_k, która ustaliła się w kalorymetrze. Wyznaczyć ciepło topnienia badanego ciała.

Ciało badane pobiera ciepło: 1) ogrzewając się jako ciało stałe do temperatury topnienia, 2) topiąc się w stałej temperaturze topnienia, 3) ogrzewając się do temperatury końcowej jako ciało ciekłe o cieple właściwym c_x''. Oddają ciepło kalorymetr i woda. Bilans cieplny ma zatem postać

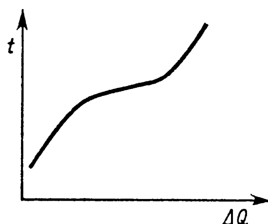
$$\Delta Q_1 + \Delta Q_2 + \Delta Q_3 = \Delta Q_4 + \Delta Q_5,$$

czyli

$$m_x c_x'(\theta - t) + m_x x + m_x c_x''(t_k - \theta) = (m_w c_w + m_k c_k)(t_0 - t_k).$$

Stąd już łatwo wyznaczyć ciepło topnienia x.

Raz jeszcze przypominamy, że wykres na rys. 17.1 dotyczy procesu topnienia ciał krystalicznych. Ciała bezpostaciowe (czyli jak pamiętamy — ciała, które można traktować jako przechłodzone ciecze) nie mają określonej temperatury topnienia. W ciałach takich podczas ogrzewania obserwuje się stopniowe mięknięcie przy ciągłym wzroście temperatury. Do takich ciał odnosi się wykres z rys. 17.2, zupełnie różny co do przebiegu od wykresu z rys. 17.1.



Rys. 17.2

17.2. Zależność temperatury topnienia od ciśnienia

Podając w poprzednim punkcie określenie temperatury topnienia podkreślaliśmy, że jest to wielkość charakterystyczna dla danego ciała pod określonym ciśnieniem. Innymi słowy, temperatura topnienia (i krzepnięcia) jest funkcją ciśnienia. W tabeli 17.1 podane są przykładowe wartości temperatur topnienia i ciepła topnienia dla różnych ciał pod ciśnieniem normalnym.

Tabela 17.1

Temperatury topnienia i ciepła topnienia pod ciśnieniem 101 325 N/m² (1 atm)

Rodzaj substancji	Temperatura topnienia °C	Ciepło topnienia · 10^3 J/kg
wodór	−257,4	62,8
tlen	−218,5	13,8
rtęć	−38,87	11,7
ołów	327,4	24,7
glin	658,8	322,0
miedź	1083	213
lód	0	334

Dla większości ciał obserwuje się wzrost temperatury topnienia ze wzrostem ciśnienia. Dla niektórych ciał są to zmiany niewielkie. Tak np. wzrost ciśnienia od 1 do 500 atm powoduje wzrost temperatury topnienia cyny o 1°C, a fosforu o około 15°C.

Do nielicznych ciał, których temperatury topnienia maleją ze wzrostem ciśnienia, należy lód, żelazo lane, bizmut, stop ołowiu, antymonu i cyny, używany do odlewania

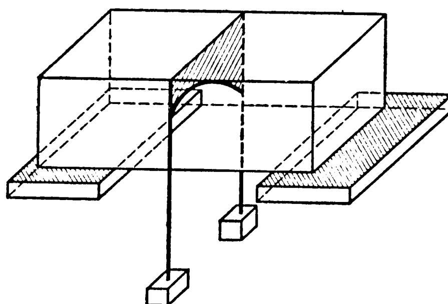
czcionek. W tabeli 17.2 podane są zmiany temperatury topnienia lodu i cyny w zależności od ciśnienia.

Tabela 17.2

Zależność temperatury topnienia od ciśnienia

Lód		Cyna	
Ciśnienie kG/cm²	Temperatura topnienia °C	Ciśnienie kG/cm²	Temperatura topnienia °C
1	0	1	231,5
336	−2,5	490	232,4
615	−5,0	1255	234,4
1155	−10	2300	236,8
2200	−22,1	3125	288,2

Różny wpływ zwiększonego ciśnienia na wysokość temperatury topnienia wiąże się ze zmianami objętości ciał podczas topnienia. Większość ciał rozszerza się podczas topnienia, a tylko nieliczne rozszerzają się podczas krzepnięcia. Do tych ostatnich należy właśnie woda i żelazo lane. Zmiany objętości wody podczas krzepnięcia są bardzo duże i dochodzą do 9%. Są to zmiany tak wielkie, że gdy woda zamarza w zamkniętych, wypełnionych po brzegi zbiornikach, to siły towarzyszące temu procesowi zdolne są rozsadzić nawet grube ścianki żelazne. Zamarzając w szczelinach skalnych woda rozsadza skały przyspieszając ich wietrzenie.



Rys. 17.3

Jeśli ciało podczas topnienia rozszerza się, to zwiększone ciśnienie jest czynnikiem utrudniającym ten proces. W tym przypadku dopiero po doprowadzeniu większej ilości energii cieplnej — po ogrzaniu do wyższej temperatury — rozpoczyna się proces topnienia.

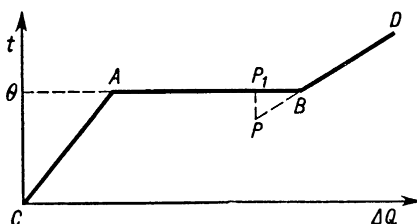
Odwrotnie, zwiększenie ciśnienia na powierzchni lodu może spowodować jego topnienie nawet wtedy, gdy otoczenie ma temperaturę 0°C. Średnio zwiększenie ciśnienia o 1 atm obniża temperaturę topnienia lodu o 0,0074 K, czyli do obniżenia temperatury topnienia lodu o 1 K trzeba powiększyć ciśnienie w przybliżeniu o 130 atm.

Przerzućmy przez bryłkę lodu o temperaturze 0°C cienki drut obciążony na obu końcach dość znacznymi ciężarkami (rys. 17.3). Wywołamy w ten sposób duże ciśnienie drutu na lód. Pod takim ciśnieniem lód się topi pobierając ciepło od swego najbliższego otoczenia, tzn. od otaczającego lodu. Drut przesuwa się ku dołowi przez warstwę wody, która następnie stykając się z lodem oziębionym nieco poniżej 0°C z powrotem zamarza. Pod drutem w dalszym ciągu odbywa się topnienie, aż wreszcie po pewnym czasie drut przechodzi przez całą bryłę. Bryła jednak nie rozpada się na dwie części dzięki procesowi regelacji, czyli ponownego skrzepnięcia wody.

Podobnie, zwiększone ciśnienie wywołuje topnienie lodu pod ostrzem łyżwy, pod płozami sań, pod powierzchnią lodowca. Ciała te przesuwają się po warstwie wody. Dzięki temu podczas ruchu sań i łyżew występuje niewielkie stosunkowo tarcie żelaza o wodę.

17.3. Przechłodzenie cieczy

Przy omawianiu krzepnięcia należy wspomnieć o tym, że niekiedy przy oziębianiu można otrzymać ciało w stanie ciekłym poniżej temperatury krzepnięcia. Osiąga się to stosując stopniowe, bardzo powolne oziębianie cieczy bez mechanicznych wstrząsów. Uzyskuje się w ten sposób *ciecz* tzw. *przechłodzoną*. Jest to stan raczej wyjątkowy, stan przekroczenia warunków równowagi w cieczy. Wystarczy do takiej przechłodzonej cieczy wrzucić kryształek ciała stałego, by wywołać gwałtowne krzepnięcie, np. tworzenie się lodu w wodzie przechłodzonej. Świadczy to o tym, że dla wystąpienia procesu krystalizacji konieczne jest *istnienie zarodków krystalizacji*. Linia kreskowana BP na rys. 17.4 odnosi się do wody przechłodzonej. W chwili gdy stan wody jest określony współrzędnymi, np. punktu P, wrzucamy do wody kawałek lodu. Będzie on stanowił właśnie zarodek krystalizacyjny. Niezwłocznie po wrzuceniu lodu rozpoczyna się gwałtowne krzepnięcie wody. Towarzyszy temu wzrost temperatury do 0°C (na wykresie odcinek PP_1). Udało się uzyskać przechłodzenie wody do -12°C.



Rys. 17.4

W doświadczeniach z naftaliną lub tiosiarczanem sodu można łatwo demonstrować występowanie stanu przechłodzenia cieczy i gwałtowny wzrost temperatury do temperatury krzepnięcia po wrzuceniu zarodka krystalizacyjnego.

Ciecze bardzo silnie przechłodzone odznaczają się tak dużą lepkością, że zachowują się podobnie jak ciała stałe. Zaliczyliśmy je do *ciał* tzw. *bezpostaciowych*.

Na zakończenie jeszcze jedna uwaga ogólna. We wstępie do tego rozdziału wspomnieliśmy, że niektóre ciała w stanie stałym występują w różnych fazach. Takimi różnymi fazami mogą być różne odmiany krystaliczne danego ciała, często w odniesieniu

do pierwiastków nazywane *odmianami alotropowymi*. Tak np. siarka, której kryształy w temperaturach bliskich temperaturze pokojowej należą do układu rombowego, przechodzi w temperaturze 95,6°C pod normalnym ciśnieniem w siarkę krystaliczną jednoskośną. Znane są też odmiany alotropowe cyny (cyna biała i szara), węgla (grafit i diament), żelaza itd. W przypadku żelaza odmiana alotropowa Fe_α ma sieć krystaliczną regularną scentrowaną przestrzennie, odmiana zaś Fe_γ — scentrowaną powierzchniowo (por. rys. 16.14).

17.4. Zjawisko parowania

Parowaniem nazywamy przejście ciała ze stanu ciekłego w stan gazowy. Możemy stworzyć sobie obraz zjawiska parowania z punktu widzenia teorii molekularno-kinetycznej. W cieczy występują cząsteczki o różnych prędkościach. Prędkościom dużym odpowiadają duże energie kinetyczne. Jeżeli cząsteczki obdarzone dużymi energiami kinetycznymi znajdą się w pobliżu powierzchni swobodnej cieczy, to mogą w swym ruchu oddalić się poza obszar przyciągania pozostałych cząsteczek cieczy. Takie cząsteczki tracą więź z cieczą — przestają być cząsteczkami cieczy, a stają się cząsteczkami pary danej cieczy.

Mówiliśmy, że średnia energia kinetyczna ruchu molekularnego jest miarą temperatury ciała. Podczas parowania ulatniają się z cieczy cząsteczki obdarzone dużymi prędkościami. Wpływa to na zmniejszenie średniej energii kinetycznej cząsteczek pozostałej cieczy. A zatem *proces parowania zachodzący samorzutnie* (bez podgrzewania cieczy) *obniża temperaturę cieczy*.

Obniżenie temperatury cieczy parującej wiąże się z oddawaniem ciepła na podtrzymanie procesu parowania.

Parowanie jest zjawiskiem odbywającym się stale na powierzchni swobodnej cieczy. Ciecz paruje w każdej temperaturze. Temperatura ma tylko wpływ na szybkość parowania. W temperaturach wyższych parowanie odbywa się szybciej: więcej jest cząsteczek o prędkościach dostatecznie dużych, aby kosztem ich energii kinetycznej mogła być wykonana praca wyrwania cząsteczki poza obręb cieczy.

Podczas stopniowego ogrzewania cieczy zwiększa się szybkość parowania. Na razie jednak przemiana cieczy w parę odbywa się w dalszym ciągu tylko na powierzchni swobodnej cieczy. Dopiero przy pewnej temperaturze charakterystycznej dla danej cieczy pod danym ciśnieniem rozpoczyna się proces parowania w całej objętości. Pęcherzyki pary tworzą się wewnątrz cieczy, stopniowo rosną, podnoszą się do góry i wydostają poza obręb cieczy.

To burzliwe parowanie nazywamy *wrzeniem cieczy*. Temperaturę, w której odbywa się wrzenie, nazywamy *temperaturą wrzenia*. Jest ona różna dla różnych cieczy. Wysokość temperatury wrzenia w dużym stopniu zależy od ciśnienia. Zwiększenie ciśnienia nad cieczą podnosi jej temperaturę wrzenia. W tab. 17.3 podane są wartości temperatur wrzenia różnych cieczy pod ciśnieniem 101 325 N/m^2 (1 atm), a tab. 17.4 ilustruje wpływ ciśnienia na wysokość temperatury wrzenia wody.

Tabela 17.3

Temperatury wrzenia i ciepła parowania pod ciśnieniem 101 325 N/m² (1 atm)

Rodzaj cieczy	Temperatura wrzenia °C	Ciepło parowania 10^3 J/kg
hel	−268,9	25,1
wodór	−252,78	452
azot	−195,81	199
tlen	−182,97	213
eter etylowy	34,6	373
alkohol etylowy	78,4	855
benzen	80,1	395
woda	100	2260
rtęć	356,9	296
siarka	444,55	
ołów	1750	
platyna	3800	

Tabela 17.4

Zależność temperatury wrzenia wody od ciśnienia

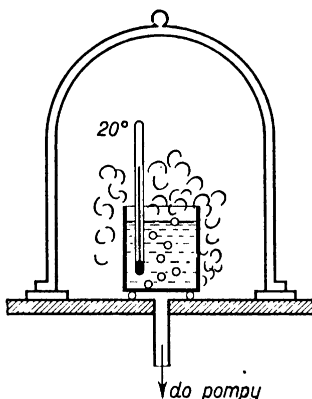
Temperatura wrzenia °C	Ciśnienie mm Hg	Temperatura wrzenia °C	Ciśnienie atm
1	4,93	120,6	2
4	6,1	152,3	5
10	9,2	180,5	10
20	17,54	198,9	15
50	92,51	224,6	25
80	355,1	374,2*	217,8*
100	760		

* Wartości odpowiadające stanowi krytycznemu wody.

Przez ogrzewanie wody w kotłach hermetycznie zamkniętych powodujemy zbieranie się pary nad powierzchnią swobodną cieczy i wzrost ciśnienia. W ten sposób można doprowadzić wodę do wrzenia w temperaturach znacznie wyższych od 100°C. Ma to zastosowanie między innymi w kotłach parowych, autoklawach, w kotłach do przerabiania kości, do przerabiania szmat na papier itd.

Można też wywołać wrzenie wody w temperaturach znacznie niższych od 100°C. Wystarczy zlewkę z wodą o temperaturze pokojowej (20°C) wstawić pod klosz pompy powietrznej (rys. 17.5) i puścić pompę w ruch. Po pewnym czasie, a mianowicie wtedy, gdy ciśnienie pod kloszem spadnie do 17,4 mm Hg, rozpocznie się wrzenie wody.

Jeśli pompa wydajnie pracuje, usuwa tworzącą się parę i w dalszym ciągu obniża ciśnienie nad cieczą, to wrzenie nie ustaje, a termometr zanurzony w wodzie wskazuje coraz niższą temperaturę. Wrzenie odbywa się kosztem ciepła pobranego z pozostałej cieczy.



Rys. 17.5

17.5. Ciepło parowania i ciepło skraplania

Parowanie cieczy odbywa się zawsze kosztem ciepła. Ilość ciepła ΔQ potrzebna do wyparowania m kg cieczy w stałej temperaturze jest proporcjonalna do masy cieczy:

$$\Delta Q = mr. \tag{17.3}$$

Współczynnik r nazywamy ciepłem parowania cieczy w danej temperaturze. Ciepło parowania wyraża się w J/kg. *Ciepło parowania* wyraża liczbowo ilość ciepła potrzebną do odparowania 1 kg danej cieczy w stałej temperaturze.

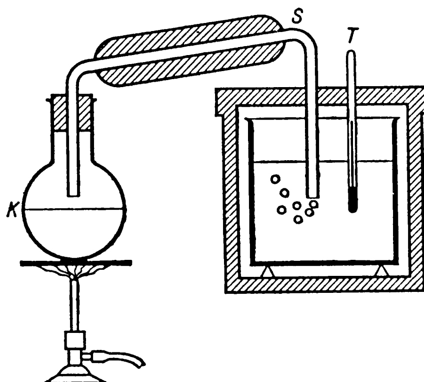
Wartości ciepła parowania zależne są od rodzaju cieczy, a dla danej cieczy zależne są od temperatury. W tab. 17.3 podano wartości ciepła parowania różnych cieczy w ich temperaturach wrzenia pod normalnym ciśnieniem, a w tab. 17.5 — zależność ciepła parowania wody od temperatury. Warto zwrócić uwagę na temperaturę, w której ciepło parowania równa się zeru. Temperaturę tę nazywamy *krytyczną*. Do tego pojęcia jeszcze wrócimy.

Tabela 17.5

Zależność ciepła parowania wody od temperatury

Temperatura °C	Ciepło parowania 10^4 J/kg
0	249
20	245
50	237
100	226
200	196
300	145
374,2	0

Podczas zamiany cieczy na parę pobierane jest ciepło z otoczenia, a w procesie odwrotnym, tzn. podczas skraplania pary, ciepło jest oddawane. *Ciepło skraplania* wyraża liczbowo ilość ciepła oddawaną przez 1 kg pary podczas skraplania w stałej temperaturze. Ciepło skraplania pary w danych warunkach jest równe ciepłu parowania cieczy w tych samych warunkach.



Rys. 17.6

Znaną z kursu szkolnego zasadę pomiaru ciepła skraplania wody w temperaturze wrzenia przypomina rys. 17.6. W kociołku K doprowadzamy wodę do wrzenia. Kociołek zamknięty jest korkiem, przez który przeprowadzona jest rurka szklana S. Dla odizolowania cieplnego od otoczenia rurka owinięta jest azbestem. Przez wylot rurki uchodzi przez kilka minut para wrzącej wody ogrzewając rurkę do temperatury wrzenia. Następnie wylot rurki wprowadzamy na kilka minut do wody w kalorymetrze. Para wodna ulega skraplaniu i oddaje ciepło. Wskutek tego ogrzewa się woda i kalorymetr. Masę skroplonej pary (m_p) wyznaczamy przez zważenie kalorymetru przed wprowadzeniem pary i po zamknięciu jej dopływu.

Zadanie. Ustalić postać bilansu cieplnego i wyrażenia na ciepło skraplania, nawiązując do doświadczenia przedstawionego na rys. 17.6 i wprowadzając następujące oznaczenia: temperatura początkowa wody w kalorymetrze t_0, końcowa t_k, masa kalorymetru m_k, wody m_w, skroplonej pary m_p, ciepło właściwe kalorymetru c_k, ciepło właściwe wody c_w.

Z dokładnych pomiarów otrzymuje się wartość ciepła skraplania (parowania) wody $r = 2257 \cdot 10^3$ J/kg dla temperatury wrzenia 100°C.

Jak widać z tab. 17.3, jest to wartość duża w porównaniu z ciepłem parowania innych cieczy. Wyparowanie wody wymaga dopływu, a skraplanie wody — dostarcza względnie dużych ilości ciepła.

Woda wyróżnia się spośród innych ciał nie tylko dużym ciepłem parowania. Jak widać z tab. 15.1 i 17.1, ma ona też stosunkowo duże ciepło właściwe i duże ciepło krzepnięcia. Duże ciepło właściwe wody decyduje o występowaniu różnicy między klimatem morskim i lądowym, przyczynia się do zmniejszenia wahań temperatury w okolicach nadmorskich. Dzięki dużemu ciepłu topnienia lodu pokrywa śnieżna utrzymuje się w zimie i nie topnieje gwałtownie na wiosnę. Duże ciepło parowania przyczynia się do powolnego parowania naturalnych zbiorników wody. Ogólnie znane jest praktyczne wykorzystanie dużych wartości ciepła właściwego i ciepła parowania wody w instalacjach centralnego ogrzewania.

17.6. Przegrzanie cieczy

Zapoznaliśmy się już poprzednio ze zjawiskiem przechłodzenia cieczy. Obecnie zajmiemy się innym zjawiskiem przekroczenia równowagi cieczy, a mianowicie *zjawiskiem przegrzania*.

Gdy ciecz przez czas dłuższy poddamy wrzeniu, a następnie oziębimy i ponownie zaczniemy bardzo powoli ogrzewać, wówczas mimo osiągnięcia temperatury wrzenia, odpowiadającej danemu ciśnieniu, ciecz może nie zacząć wrzeć. Jako skutek ogrzewania wystąpi dalszy wzrost temperatury cieczy. Otrzymamy ciecz tzw. przegrzaną. Wprowadzenie do takiej cieczy pęcherzyków gazu wywołuje gwałtowne parowanie i spadek temperatury do wartości temperatury wrzenia. Na parowanie cieczy zostaje zużyte ciepło pobrane poprzednio podczas ogrzewania się cieczy ponad temperaturę wrzenia.

Ze względu na wydzielanie się ciepła w całej objętości, parowanie cieczy przegrzanej przebiega zazwyczaj bardzo gwałtownie. Jeśli ten proces odbywa się w naczyniu zamkniętym, towarzyszący mu wzrost ciśnienia wywołany powstaniem dużej ilości pary może nawet doprowadzić do rozsadzenia naczynia.

Jak zaznaczyliśmy, parowanie odbywa się na powierzchni swobodnej cieczy. Na ściankach naczynia lub na powierzchni ciał stałych zanurzonych w cieczy znajdują się zwykle mikroskopijnej wielkości pęcherzyki zaadsorbowanych gazów. Powierzchnie tych pęcherzyków stanowią również powierzchnię swobodną cieczy. Jeśli temperatura cieczy równa się temperaturze wrzenia, to na powierzchni tych pęcherzyków odbywa się gwałtowne parowanie. Pęcherzyk rozszerza się i pod wpływem wyporu wypływa na powierzchnię. Rozpoczyna się proces wrzenia. Do wrzenia cieczy niezbędne jest istnienie w jej wnętrzu pęcherzyków gazu jako zarodków tworzącej się pary. Dłuższe gotowanie usuwa z cieczy pęcherzyki gazów i umożliwia uzyskanie stanu przekroczenia równowagi, jakim jest stan cieczy przegrzanej.

17.7. Para nasycona i nienasycona

Zajmiemy się obecnie zjawiskiem parowania cieczy w naczyniach zamkniętych.

Zbiornik A (rys. 17.7) połączony jest z manometrem rtęciowym. Notujemy stan początkowy manometru. Przez otwieranie kurka K wpuszczamy kroplami do naczynia ciecz lotną, tzn. szybko parującą już w temperaturze pokojowej, np. eter. Pierwsze krople eteru w całości wyparowują, manometr wykazuje zwyżkę ciśnienia w zbiorniku. Do gazów istniejących poprzednio w zbiorniku dołącza się nowy składnik — para eteru — wywierający pewne ciśnienie cząstkowe i przyczyniający się zgodnie z prawem Daltona do wzrostu ciśnienia całej mieszaniny.

W miarę wpuszczania większej ilości cieczy zwiększa się ilość pary zawartej w naczyniu, rośnie gęstość pary i prężność.

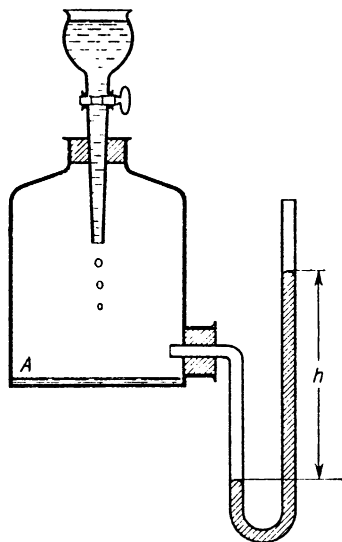
Ten stan rzeczy trwa do chwili, gdy pewna z kolei wpuszczona porcja eteru już nie wyparowuje, lecz pozostaje w naczyniu w postaci cieczy. W tej chwili manometr wskazuje maksymalną zwyżkę ciśnienia w stosunku do stanu początkowego. Zwyżka ta stanowi największe ciśnienie pary eteru w danej temperaturze, nie ulegające zmianie pod-

czas wpuszczania dalszych kropli cieczy. W objętości zbiornika zebrana jest maksymalna ilość pary — odpowiada temu osiągnięcie maksymalnej gęstości pary w danej temperaturze.

Parę o maksymalnej prężności, a równocześnie o maksymalnej gęstości w danej temperaturze nazywamy parą nasyconą danej cieczy. Mniejszym prężnościom pary w zbiorniku i mniejszym gęstościom odpowiada stan *pary nienasyconej.*

Z punktu widzenia makroskopowego osiągnięcie stanu nasycenia odpowiada przerwaniu procesu parowania cieczy.

Z punktu widzenia teorii kinetyczno-molekularnej stale odbywa się ucieczka najszybszych cząsteczek z powierzchni swobodnej cieczy, czyli stale odbywa się proces parowania. Równocześnie jednak zachodzi proces odwrotny: proces skraplania. Cząsteczki pary w swoim ruchu molekularnym trafiają z powrotem w sferę przyciągania cząsteczek cieczy i tam pozostają. Dopóki stan rzeczy jest taki, że istnieje przewaga po stronie liczby cząsteczek uchodzących z cieczy, dopóty nad cieczą znajduje się para nienasycona. W pewnej chwili zostaje osiągnięty *stan równowagi statystycznej*: w danym czasie tyle cząsteczek przechodzi ze stanu ciekłego w stan pary, ile cząsteczek przechodzi ze stanu pary w stan ciekły (tzn. tyle paruje, ile równocześnie się skrapla). W naczyniu istnieje para nasycona w danej temperaturze. Mówimy, że *para nasycona pozostaje w równowadze z cieczą.*



Rys. 17.7

Warunek równowagi statystycznej można ująć matematycznie jako

$$\frac{dm}{dt} = 0$$

w stałej temperaturze.

Innymi słowy, pochodna masy (czy to pary nasyconej, czy też cieczy pozostającej z nią w zetknięciu) względem czasu równa się zeru. Jest to równoznaczne z niezmiennością masy cieczy lub masy pary nasyconej w zamkniętym naczyniu w stałej temperaturze.

Z tych rozważań wynika, że parę nasyconą danej cieczy możemy w każdej temperaturze charakteryzować przez podanie jej prężności lub gęstości. W tab. 17.6 podano kilka wartości prężności par nasyconych w rozmaitych temperaturach.

W tabeli 17.6 zwracają uwagę powtarzające się wartości prężności pary nasyconej równe 760 mm Hg. Taką prężność ma para wodna przy 100°C, para eteru przy 34,6°C, para alkoholu przy 78,4°C, para rtęci przy 356,9°C itd.

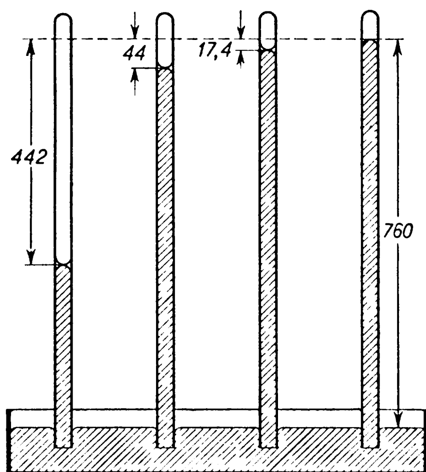
Wymienione temperatury są temperaturami wrzenia tych cieczy pod ciśnieniem normalnym.

Innymi słowy, *temperatura wrzenia jest to taka temperatura, w której prężność pary nasyconej danej cieczy równa się ciśnieniu zewnętrznemu.*

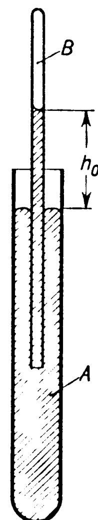
Tabela 17.6

Prężności pary nasyconej różnych cieczy

Rodzaj cieczy	Temperatura °C	Prężność pary nasyconej mm Hg
eter etylowy	−20	63
	20	442
	34,6	760
alkohol etylowy	0	12,7
	30	78,4
	78,4	760
	100	1698
rtęć	0	0,0002
	20	0,0013
	100	0,270
	356,9	760
	500	6080, czyli 8 atm
woda	0	4,579
	20	17,54
	100	760
	300	64450, czyli ok. 85 atm



Rys. 17.8



Rys. 17.9

Łatwo można wykazać doświadczalnie różnice prężności par nasyconych różnych cieczy w określonej temperaturze. W szerokim płaskim naczyniu z rtęcią (rys. 17.8) zanurzone są 4 barometryczne rurki szklane o wysokości ok. 1 metra. Do rurki pierwszej wprowadzona jest taka ilość eteru, że na powierzchni rtęci mamy cienką warstwę ciekłego eteru. Świadczy to o tym, że w przestrzeni nad cieczą w rurce pierwszej istnieje para nasycona rtęci i etéru. Podobnie w rurce drugiej mamy cienką warstwę ciekłego

alkoholu, w rurce trzeciej cienką warstwę wody. W rurce czwartej mamy tylko rtęć i nad nią próżnię Torricellego. Poziomy powierzchni swobodnych rtęci we wszystkich rurkach są niejednakowe, co świadczy o różnych ciśnieniach wywieranych na powierzchnie. W rurce czwartej ciśnienie wywierane jest tylko przez parę nasyconą rtęci. W rurkach pozostałych prócz ciśnienia pary rtęci mamy dodatkowo ciśnienia par nasyconych odpowiednich cieczy: eteru, alkoholu lub wody (ciśnienie wywołane ciężarem cienkiej warstwy cieczy zaniedbujemy). Prężność pary nasyconej eteru, alkoholu i wody wywołuje obniżenie poziomu rtęci odpowiednio o 442 mm, 44 mm i 17,4 mm. Liczby te odnoszą się do pomiaru przeprowadzonego w temperaturze 20°C.

Można też doświadczalnie sprawdzić, że prężność pary nasyconej nie zależy od objętości. W naczyniu *A* (rys. 17.9) wypełnionym rtęcią można dowolnie głęboko zanurzać rurkę szklaną *B*, zamkniętą u góry. Do rurki wprowadzona jest taka ilość eteru, że na powierzchni rtęci znajduje się warstewka ciekłego eteru. Zacznijmy podnosić rurkę *B* stopniowo do góry. Różnica poziomów rtęci w naczyniach *A* i *B* pozostaje niezmienna. Zwiększa się objętość nad cieczą wypełniona parą nasyconą i zmniejsza się grubość warstewki ciekłego eteru.

Ciśnienie atmosferyczne p_0, działające na powierzchnię rtęci w naczyniu *A*, jest w dalszym ciągu równoważone przez ciśnienie hydrostatyczne słupa rtęci o wysokości h_0, ciśnienie warstwy eteru, ciśnienie pary nasyconej rtęci i ciśnienie pary nasyconej eteru p_{et}. Z wymienionych ciśnień możemy zaniedbać jako bardzo małe ciśnienie warstwy eteru i ciśnienie pary nasyconej rtęci. A więc

$$p_0 = h_0 \varrho g + p_{et}.$$

Ponieważ podczas przesuwania rurki *B* wysokość h_0 nie ulega zmianie, przeto ze stałości p_0 i $h_0 \varrho g$ wynika niezależność prężności pary nasyconej eteru od objętości.

Podczas zagłębiania rurki *B* różnica poziomów rtęci w naczyniach *A* i *B* pozostaje również niezmienna. Zmniejszaniu objętości nad cieczą w rurce *B* towarzyszy skraplanie pary, gdyż warstewka ciekłego eteru staje się coraz grubsza.

Wynik doświadczenia uogólniamy na wszystkie ciecze: *prężność pary nasyconej nie zależy od objętości. Podczas zwiększania objętości pewna masa cieczy dodatkowo wyparowuje, podczas zmniejszania objętości — nadmiar pary się skrapla.* Gęstość pary nasyconej w stałej temperaturze pozostaje niezmienna.

Spróbujmy jednak w tym samym doświadczeniu rurkę *B* wyciągać coraz bardziej. Mówiliśmy, że grubość warstewki eteru stopniowo maleje, wreszcie przy pewnym położeniu rurki *B* warstewka ciekłego eteru znika. Jeszcze w tym momencie utrzymuje się nad rtęcią para nasycona eteru, zwana *parą nasyconą suchą*. Różnica poziomów rtęci w naczyniu *A* i *B* pozostaje jeszcze niezmieniona. Wystarczy jednak najmniejsze dalsze przesunięcie rurki *B* w górę, aby w związku ze wzrostem objętości para nasycona sucha przekształciła się na parę nienasyconą. Niezwłocznie też podnosi się poziom rtęci w rurce *B*, tzn. rośnie h_0, świadcząc o *zmniejszaniu się prężności pary nienasyconej podczas wzrostu objętości*. Doświadczenia te należy przeprowadzać w stałej temperaturze. Im bardziej w danej temperaturze zwiększamy objętość pary nienasyconej, tym bardziej staje się ona odległa od stanu nasycenia. Mierząc ciśnienia pary nienasyconej

i odpowiadające im objętości można stwierdzić, że do par nienasyconych dalekich od stanu nasycenia stosuje się prawo Boyle'a z tym samym przybliżeniem, jak do gazów rzeczywistych.

Doświadczenia wykazują, że do pary nienasyconej, dalekiej od stanu nasycenia, stosują się też z tym samym przybliżeniem prawa Charles'a–Gay-Lussaca. Innymi słowy, *pary nienasycone dalekie od stanu nasycenia podlegają tym samym prawom, co gazy rzeczywiste.*

Nie można tego powiedzieć o parach nasyconych. Stwierdziliśmy już, że nie stosuje się do nich prawo Boyle'a, gdyż w stałej temperaturze prężność pary jest niezależna od objętości. Nie stosują się też do nich pozostałe prawa gazowe.

Prężność pary nasyconej ogrzewanej w stałej objętości wzrasta gwałtowniej niż ciśnienie gazu ogrzewanego w tych samych warunkach.

Łatwo to wyjaśnimy. Z punktu widzenia makroskopowego wzrost ciśnienia gazu pod wpływem ogrzania jest wynikiem tylko ograniczenia możliwości rozszerzania się. Wzrost prężności pary nasyconej jest spowodowany również ograniczeniem możliwości rozszerzania się, ponadto jednak wchodzi w grę drugi bardzo ważny czynnik: przy wzroście temperatury odbywa się dodatkowe parowanie cieczy, czyli zwiększa się ilość pary naciskającej na ścianki. Te dwa czynniki wywołują szybkie wzrastanie prężności pary nasyconej w miarę podnoszenia temperatury.

Jeśli w naczyniu zamkniętym znajduje się niewielka ilość cieczy, to przy stopniowym podnoszeniu temperatury obserwujemy zmniejszanie się ilości cieczy i wreszcie zupełny jej zanik.

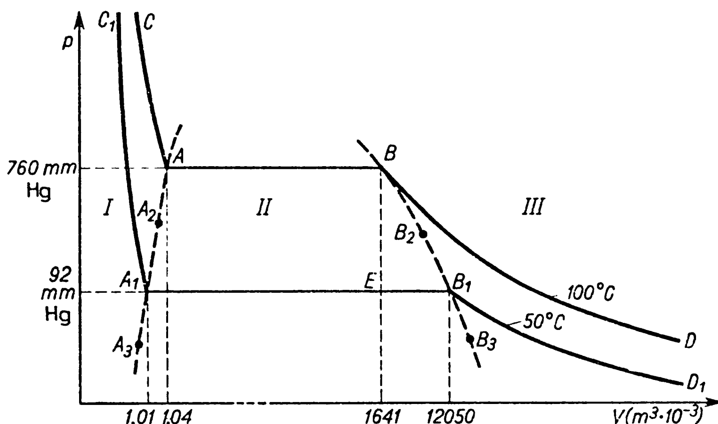
W momencie wyparowania ostatniej kropli cieczy w naczyniu jest jeszcze para nasycona sucha. Najmniejsze podgrzanie przekształca ją w parę nienasyconą. Tak więc *parę nasyconą suchą można przekształcić na parę nienasyconą nie tylko przez zwiększenie izotermiczne objętości, lecz również przez izochoryczne lub izobaryczne ogrzanie.* Dlatego też parę nienasyconą nazywa się często parą przegrzaną.

17.8. Graficzne przedstawienie właściwości par

Ogólnie przyjętym sposobem przedstawiania właściwości par jest kreślenie krzywych zależności ciśnienia od objętości w stałych temperaturach. Krzywe te noszą nazwę izoterm.

Aby lepiej poznać sposób kreślenia i sens fizyczny otrzymywanych krzywych, rozważmy początkowo zachowanie się 1 kg wody utrzymywanego w stałej temperaturze i zamkniętego w naczyniu o szczelnie dopasowanym tłoku. Na osi odciętych odkładamy objętości, na osi rzędnych — ciśnienia (rys. 17.10). Przeanalizujmy przebieg górnej krzywej wykresu odnoszącej się do temperatury 100°C. Dopóki tłok jest silnie obciążony, ciśnienie jest większe od 760 mm Hg (1 atm), woda występuje w stanie ciekłym. Objętość jej mimo dużego ciśnienia mało się różni od wartości $1{,}04 \cdot 10^{-3}$ m^3 odpowiadającej objętości 1 kg wody w temperaturze 100°C pod ciśnieniem normalnym.

Wynika to z małej ściśliwości cieczy. Zmniejszając ciśnienie na tłok powodujemy minimalne zwiększanie się objętości wody. Odpowiada to na wykresie stromo opadającej linii *CA*. Punkt *A* o współrzędnych $V_A = 1{,}04 \cdot 10^{-3}$ m³ i $p_A = 760$ mm Hg określa taki stan, kiedy jeszcze pod tłokiem istnieje tylko ciecz, a nie ma pary. Wystarczy jednak, doprowadzając ciepło, choć trochę powiększyć objętość, aby dzięki procesowi



Rys. 17.10

wrzenia znalazła się pod tłokiem obok cieczy jej para nasycona. Ciśnienie przy tym nie ulega zmianie. Odcinek prostoliniowy sięgający od *A* do *B* odpowiada równoczesnemu istnieniu pod tłokiem (współistnieniu) cieczy i jej pary nasyconej. Punkt *B* o współrzędnych $V_B = 1{,}641$ m³ i $p_B = 760$ mm Hg (skala na wykresie nie jest zachowana) określa występowanie pary nasyconej suchej. Dalsze zwiększenie objętości powoduje przekształcenie się pary nasyconej suchej na parę nienasyconą. Ciśnienie pary nienasyconej maleje ze wzrostem objętości. Na wykresie przedstawia to linia *BD*, przechodząca w hiperbolę w części dalszej, odnoszącej się do pary nienasyconej odległej od stanu nasycenia.

Analogiczny przebieg ma izoterma dolna, odnosząca się do temperatury 50°C. Widzimy na niej stromy przebieg linii C_1A_1, wyrażający małą ściśliwość cieczy. Punkt A_1 o współrzędnych $V_{A_1} = 1{,}01 \cdot 10^{-3}$ m³, $p_{A_1} = 92$ mm Hg odpowiada ostatniemu momentowi występowania pod tłokiem samej tylko cieczy. Odcinek od A_1 do B_1 określa współistnienie cieczy i pary nasyconej. Punkt B_1 o współrzędnych $V_{B_1} = 12{,}050$ m³ i $p_{B_1} = 92$ mm Hg odnosi się do pary nasyconej suchej. Krzywa B_1D_1 ilustruje własności pary nienasyconej w temperaturze 50°C.

Zaznaczyliśmy na wykresie tylko dwie izotermy, odpowiadające temperaturom 100 i 50°C. Moglibyśmy ich nakreślić bardzo wiele, przechodząc stopniowo od temperatur niskich do coraz wyższych. W przedziale 50–100°C na każdej z takich izoterm wystąpiłby odcinek odpowiadający cieczy, taki jak *CA*, odcinek współistnienia cieczy

i pary nasyconej, jak *AB*, i linia krzywa typu *BD*, odpowiadająca parze nienasyconej. Połączenie punktów A, A_1, A_2, A_3 ... oraz B, B_1, B_2, B_3 ... zaznaczone na rysunku liniami kreskowanymi wydziela na wykresie trzy obszary:

I — obszar cieczy,

II — obszar współistnienia cieczy i pary,

III — obszar pary nienasyconej.

Jak widać z wykresu, w miarę przechodzenia do temperatur wyższych obszar współistnienia cieczy i pary coraz bardziej się zwęża, gdyż odcinek *AB* staje się coraz krótszy.

Z wykresu można odczytać i inne fakty. Tak np. można sprawdzić, jakie są możliwości skroplenia pary nasyconej suchej. Stan pary nasyconej suchej określają na wykresie punkty typu *B*. Dla skroplenia wystarczy:

1) przesunąć się wzdłuż *BA*, tzn. izotermicznie zmniejszyć objętość pary nasyconej suchej (będzie to równocześnie przemiana izobaryczna, gdyż zmiana objętości pary nasyconej nie wpływa na jej prężność), albo

2) przesunąć się wzdłuż linii *BE*, tzn. izochorycznie obniżyć temperaturę. Punkty leżące poniżej *B* (jak np. *E*) znajdują się na odcinkach współistnienia cieczy i pary izoterm odpowiadających temperaturze niższej od 100°C, a więc po osiągnięciu takiego stanu część pary jest już skroplona.

Z wykresu można wyciągnąć inny jeszcze wniosek, a mianowicie: przed skropleniem para nienasycona musi dojść do stanu nasycenia. W tym celu można zastosować np. zgęszczanie izotermiczne, oziębianie izochoryczne lub oziębianie izobaryczne.

Skroplenie pary zachodzi niejednokrotnie z pewnym opóźnieniem, tzn. mimo osiągnięcia takich warunków, że prężność pary równa się, a nawet staje się większa od prężności pary nasyconej w danej temperaturze — proces skraplania nie rozpoczyna się. Mówimy wtedy o występowaniu *pary przesyconej*. Przyczyną takiego zjawiska jest brak ośrodków kondensacji pary w postaci czy to zawiesin, pyłków, czy też jonów wytworzonych działaniem promieni Röntgena lub nadfioletu, lub też dzięki przelatującym przez parę rozpędzonym cząstkom (por. działanie komory Wilsona, § 30.2).

Skraplanie pary wodnej lub jej zestalanie się w atmosferze ziemskiej wiąże się z takimi zjawiskami, jak: deszcz, mgła, śnieg, grad, rosa, szron, sadź itp.

Mgła powstaje najczęściej przy mieszaniu się ciepłego powietrza z chłodniejszym (np. w pobliżu jezior, rzek, bagien) i polega na skraplaniu się pary wodnej w postaci małych kropelek na drobnych zawiesinach (np. kurzu lub pyłu węglowego w pobliżu ośrodków przemysłowych). Mgła osiadająca na ciałach stałych tworzy albo rosę, jeśli temperatura ciał jest wyższa od zera, albo szron.

Mgły wytwarzające się w górnych warstwach atmosfery, w tzw. troposferze na wysokości 8–18 km ponad Ziemią, tworzą chmury.

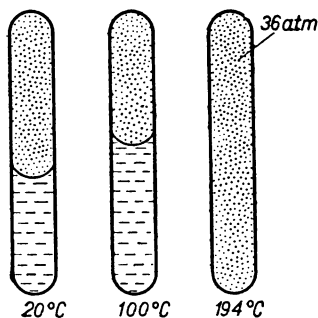
Deszcz powstaje wtedy, gdy drobne kropelki mgły w chmurach łączą się w większe krople. Gdy zjawisko zachodzi w dostatecznie niskiej temperaturze, może tworzyć się śnieg. Jeżeli natomiast w atmosferze powstają warunki odpowiednie do występowania zjawiska przechłodzenia kropel wody, to z nich może wytwarzać się grad. Przechłodzone kropelki mgły stykając się z przedmiotami stałymi na powierzchni Ziemi powodują powstawanie sadzi.

17.9. Stan krytyczny

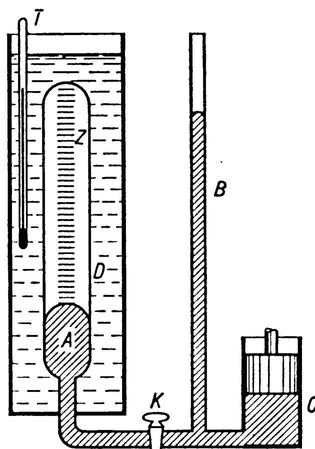
Wykres z rys. 17.10 powinien nam nasunąć do rozważenia nowe zagadnienie. Jaki będzie przebieg izoterm w temperaturach coraz wyższych? Czy nawet w dowolnie wysokich temperaturach możliwe jest współistnienie każdej substancji w stanie ciekłym i w stanie pary, innymi słowy — czy w każdej temperaturze wystąpi na wykresie odpowiedniej izotermy odcinek typu AB?

Nawiążemy do klasycznego doświadczenia Cagniarda de la Toura. W grubościennej szklanej probówce podtrzymujemy przez czas dłuższy wrzenie eteru celem usunięcia powietrza oraz rozpuszczonych i zaadsorbowanych na ściankach gazów. Następnie probówkę zatapiamy pozostawiając w niej odpowiednią ilość ciekłego eteru i parę nasyconą eteru. Obie te postaci eteru są wyraźnie rozgraniczone powierzchnią swobodną cieczy.

Zaczynamy probówkę stopniowo, ostrożnie ogrzewać. Towarzyszy temu rozszerzanie się cieczy, dalsze jej parowanie i zgęszczanie się pary nasyconej. Gęstość cieczy stopniowo maleje, gęstość pary rośnie. Powierzchnia swobodna cieczy staje się coraz bardziej płaska. Gdy podczas ogrzewania dochodzimy do temperatury 194°C, zanika powierzchnia swobodna, rozdzielająca dotychczas eter ciekły (rys. 17.11) i eter w postaci gazowej. Cała probówka jest wypełniona jednolitą substancją. W wymienionej temperaturze zanikła różnica między cieczą i parą danej substancji: wyrównały się gęstości (wszak ciecz stawała się podczas ogrzewania coraz rzadsza, a para coraz gęstsza), stały się identyczne własności optyczne. Gdybyśmy mogli w tej temperaturze zmierzyć panujące w probówce ciśnienie, otrzymalibyśmy wartość około 36 atm. Mając na uwadze tę dużą wartość ciśnienia podkreśliliśmy konieczność stosowania w tym doświadczeniu grubościennych, wytrzymałych rurek.



Rys. 17.11



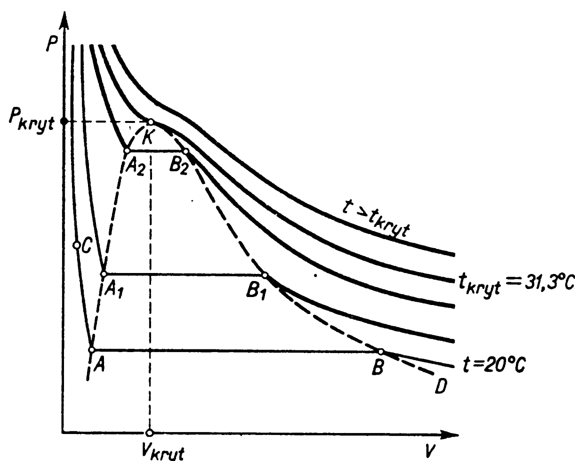
Rys. 17.12

Gdybyśmy zawartość probówki ogrzali nieco powyżej temperatury 194°C, a następnie pozwolili jej powoli się oziębiać, to znów bylibyśmy świadkami pojawienia się w temperaturze 194°C granicy rozdzielającej substancję w stanie ciekłym i w stanie

gazowym. Dalsze przemiany postępowałyby tak jak poprzednio, tylko w odwrotnej kolejności: od małych gęstości cieczy i stosunkowo dużych objętości przechodzilibyśmy do gęstości coraz większych, a objętości coraz mniejszych.

Temperatura 194°C w przypadku eteru jest temperaturą, w której zanika różnica między cieczą i parą. Temperaturę tę nazywamy *temperaturą krytyczną*. Odkrycie temperatury krytycznej miało duże znaczenie. Pchnęło na nowe tory technikę skraplania gazów przyczyniając się do opracowania metod uzyskiwania bardzo niskich temperatur. Od końca ubiegłego stulecia datuje się powstawanie i rozwój laboratoriów kriogenicznych, tzn. laboratoriów nastawionych na prowadzenie prac badawczych w temperaturach bardzo niskich, bliskich zera bezwzględnego.

Pójdziemy obecnie śladami Andrewsa (1869 r.) i prześledzimy przemiany izotermiczne dwutlenku węgla. Do badań może służyć urządzenie przedstawione na rys. 17.12. Widzimy tam naczynie połączone o trzech ramionach *Z*, *B* i *C*, przy czym część pozioma zaopatrzona jest w kurek *K*. W ramieniu *Z* znajduje się pewna ilość gazu (CO_2) zamknięta od dołu powierzchnią rtęci. Naczynie to jest zaopatrzone w podziałkę, na której bezpośrednio odczytujemy każdorazową objętość gazu. Część *Z* urządzenia zanurzona jest w termostacie *D* umożliwiającym utrzymanie stałej temperatury badania. Za pomocą pompy *C* wpychamy rtęć do ramion *A* i *B* powodując zmniejszenie objętości i równocześnie wzrost ciśnienia gazu w ramieniu *Z*. Wartość tego ciśnienia odczytujemy wykorzystując ramię *B* jako ramię manometru. Odpowiadające sobie wartości ciśnienia i objętości gazu w stałej temperaturze odkładamy na wykresie i otrzymujemy izotermy dwutlenku węgla (rys. 17.13).



Rys. 17.13

Przebieg zjawiska, a zatem i przebieg izoterm zależny jest od temperatury, w jakiej przeprowadzamy badania. Prześledźmy zjawisko i odpowiadającą mu izotermę w temperaturze 20°C. Początkowo w miarę zmniejszania objętości rośnie ciśnienie, co na wykresie odpowiada przesuwaniu się od *D* do *B*. W pewnej chwili rozpoczyna się w zbiorniku kondensacja pary — pojawia się zamglenie, potem na ściankach naczynia

tworzą się kropelki cieczy i stopniowo na powierzchni rtęci tworzy się warstwa ciekłego dwutlenku węgla. W tym czasie ciśnienie nie ulega zmianie. Na wykresie odpowiada to odcinkowi *BA* równoległemu do osi objętości. Gdy cały dwutlenek węgla zawarty w ramieniu *Z* ulegnie skropleniu, dalsze zmniejszanie objętości wymaga bardzo gwałtownego wzrostu ciśnienia — odpowiada temu część *CA* wykresu.

Omówiona krzywa ma charakter analogiczny do izoterm wody z rys. 17.10. Odcinek *AB* odpowiada współistnieniu dwutlenku węgla w postaci ciekłej i w postaci pary nasyconej. Odcinek *CA* odnosi się do ciekłego CO_2. Punkt *B* odpowiada parze nasyconej suchej. Krzywa *BD* charakteryzuje właściwości CO_2 w postaci pary nienasyconej.

W temperaturach wyższych, aż do 31,3°C, przebieg zjawiska i przebieg izoterm jest podobny. Stopniowo tylko odcinek prostoliniowy staje się coraz krótszy. Punkty typu *A* (a więc A_1, A_2 ...) przesuwają się w stronę objętości większych, gdyż objętość danej masy cieczy CO_2 w wyższej temperaturze jest nieco większa. Punkty typu *B* (B_1, B_2, ...) odpowiadają parze nasyconej suchej w coraz wyższej temperaturze, a więc parze o coraz większej gęstości. Większa gęstość przy danej masie idzie w parze z mniejszą objętością, dlatego kolejne odcinki *AB*, A_1B_1, A_2B_2 itd. stają się w miarę wzrostu temperatury coraz krótsze. Gęstość pary nasyconej suchej i ciekłego dwutlenku węgla dążą do wyrównania się.

W temperaturze 31,3°C izoterma dwutlenku węgla ma przebieg inny. Nie występuje już odcinek prostoliniowy typu *AB*, a zamiast niego mamy punkt *K* przegięcia krzywej. Na tej izotermie nie potrafimy już rozgraniczyć części odnoszącej się do dwutlenku węgla w postaci ciekłej i w postaci gazowej. W temperaturze 31,3°C nastąpiło wyrównanie właściwości substancji w stanie ciekłym i w stanie gazowym. Jest to zatem temperatura krytyczna dwutlenku węgla. Współrzędne punktu przegięcia izotermy krytycznej nazywamy zwykle objętością krytyczną V_k i ciśnieniem krytycznym p_k. Dla dwutlenku węgla temperatura krytyczna $t_k = 31{,}3°C$, $p_k = 73$ atm. Objętość krytyczna V_k zależna jest od badanej masy dwutlenku węgla.

W miarę przechodzenia do temperatur wyższych od krytycznej przebieg izoterm staje się coraz bardziej zbliżony do hiperbolicznego. W tych temperaturach mamy do czynienia z jednolitą substancją gazową.

Doświadczenie uczy, że wnioski wypływające z badań dotyczących dwutlenku węgla odnoszą się do wszystkich gazów. Każdy gaz ma określoną temperaturę krytyczną. W temperaturze wyższej od krytycznej substancje mogą występować tylko w stanie gazowym. Poniżej temperatury krytycznej można każdą substancję otrzymać w stanie ciekłym przez zastosowanie odpowiedniego ciśnienia. Innymi słowy, każdy *gaz można skroplić po oziębieniu go do temperatury niższej od temperatury krytycznej.*

17.10. Skraplanie gazów

Badania Andrewsa zmieniły pogląd na naturę gazów. Od czasów Faradaya (ok. 1820 r.) datują się zakreślone na szerszą skalę próby skraplania gazów, uwieńczone zresztą w licznych przypadkach dobrym skutkiem. Doświadczenia przeprowadzane były w ten sposób, że stosowano niezbyt wielkie oziębianie, lecz bardzo duże ciśnienia.

Mimo osiągnięcia ciśnień rzędu kilku tysięcy atmosfer (do 3000 atm) nie udało się niektórych gazów uzyskać w postaci ciekłej. Nazwano te gazy trwałymi. Do tej kategorii zaliczono tlen, azot, wodór, tlenek węgla, metan. Do tej samej grupy zaliczony byłby hel, gdyby już w okresie tym był znany na Ziemi (hel został odkryty na Ziemi w r. 1895).

Badania Andrewsa rzuciły inne światło na wyniki prac nad skraplaniem gazów. Tylko te gazy udało się skroplić, które miały temperaturę krytyczną wyższą od temperatury, w której były przeprowadzane badania. Gazy „trwałe" nie mogły ulec skropleniu w temperaturach bliskich 0°C, gdyż ich temperatury krytyczne są znacznie niższe (tab. 17.7). W stosunku do tych gazów trzeba było zastosować wstępne oziębienie, a dopiero wtedy zwiększenie ciśnienia mogło wywołać skroplenie.

Tabela 17.7

Temperatury i ciśnienia krytyczne

Rodzaj ciała	Temperatura krytyczna °C	Ciśnienie krytyczne atm
hel	−267,9	2,26
wodór	−239,9	12,8
azot	−147,1	33,5
tlen	−118,8	49,7
dwutlenek węgla	31,35	72,9
eter etylowy	193,6	36,3
alkohol etylowy	245,1	62,96
woda	374,2	217,7

Według dzisiejszych poglądów stan gazowy nie jest stanem charakterystycznym dla pewnych substancji. Gazy traktujemy jako pary nienasycone odpowiadających im cieczy. Pary te zależnie od warunków mogą być bliższe lub dalsze od stanu nasycenia.

Aby lepiej zdać sobie sprawę z możliwości i sposobów skraplania gazów, przyjrzyjmy się wykresowi (rys. 17.14). Zaznaczone są na nim cztery obszary.

Obszar pierwszy — to *obszar cieczy* między częścią izotermy krytycznej MK i linią cieczy przechodzącą przez punkty A_1, A_2, A_3 ...

Obszar drugi — *obszar współistnienia cieczy i pary nasyconej*, ograniczony linią cieczy i linią odpowiadającą stanowi pary nasyconej suchej, przechodzącą przez punkty B_1, B_2, B_3, ...

W obszarze trzecim i czwartym dana substancja występuje w postaci lotnej. Można by więc ujmować je łącznie. Wyróżnimy jednak: obszar trzeci *jako obszar pary nienasyconej o temperaturze niższej od krytycznej*. Granicami tego obszaru są linie pary nasyconej suchej i część KN izotermy krytycznej.

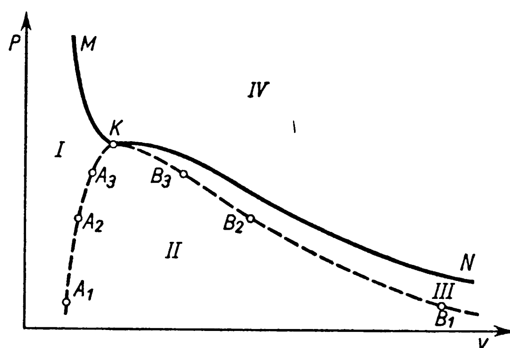
Obszar czwarty nazywać będziemy *obszarem gazu*. Obejmować on będzie wszystkie

punkty wykresu leżące powyżej izotermy krytycznej, czyli odnosi się do stanów pary nienasyconej, bardziej odległych od nasycenia, niż to ma miejsce w obszarze trzecim.

Skroplenie pary odpowiada przejściu z obszaru trzeciego lub czwartego do obszaru drugiego. Skroplenie gazu, czyli przejście z obszaru IV do II wymaga zawsze przekroczenia izotermy krytycznej, czyli obniżenia temperatury poniżej temperatury krytycznej.

Osiągnięcie temperatur znacznie niższych od 0°C nie jest sprawą łatwą. Już Faraday posługiwał się mieszaninami mrożącymi, ale na tej drodze nie można było dojść do temperatur bardzo niskich. Tak np. mieszanina mrożąca lodu z solą daje maksymalne obniżenie temperatury do —21,5°C, mieszanina lodu z chlorkiem wapnia — do około —50°C.

W technice skraplania gazów wykorzystano inne, bardziej efektywne sposoby oziębiania. Omówimy metodę kaskadową oraz metodę Hampsona i Lindego.



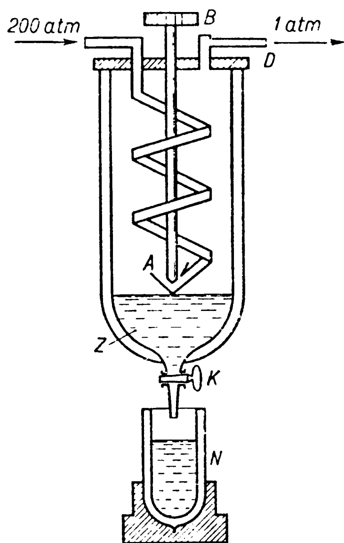
Rys. 17.14

Metoda kaskadowa polega na kolejnym skraplaniu gazów o coraz niższych temperaturach krytycznych. Pierwszy gaz o stosunkowo wysokiej temperaturze krytycznej zostaje skroplony pod ciśnieniem zwiększonym. Po skropleniu zostaje on doprowadzony do wrzenia przez zmniejszenie ciśnienia nad jego powierzchnią. Wrzenie odbywa się kosztem ciepła pobieranego z najbliższego otoczenia, a więc przede wszystkim z reszty skroplonego gazu. Dzięki temu ta ciecz się oziębia i może służyć do obniżenia temperatury gazu drugiego o niższej temperaturze krytycznej. Drugi gaz, oziębiony poniżej temperatury krytycznej, podlega temu samemu procesowi: pod zwiększonym ciśnieniem zostaje skroplony, następnie pod zmniejszonym ciśnieniem wrze i oziębia się. Po oziębieniu może służyć dalej do obniżenia temperatury gazu trzeciego o jeszcze niższej temperaturze krytycznej. Metoda ta zaproponowana przez Picteta została z dodatnim wynikiem zastosowana przez uczonych polskich Wróblewskiego i Olszewskiego, którzy w roku 1883 w laboratorium Uniwersytetu Jagiellońskiego skroplili powietrze.

Pomocniczym gazem wykorzystanym przez Wróblewskiego i Olszewskiego był etylen C_2H_4 o temperaturze krytycznej +9,5°C. Etylen oziębiony w mieszaninie lodu z solą i poddany ciśnieniu kilkudziesięciu atmosfer ulega skropleniu. Gdyby skroplony etylen był umieszczony w zbiorniku otwartym (a więc pod ciśnieniem 1 atm), to dzięki gwałtownemu parowaniu oziębiłby się do —103,5°C, tzn. do temperatury wrzenia ety-

lenu pod ciśnieniem normalnym. Ta temperatura jest jeszcze wyższa od temperatury krytycznej głównych składników powietrza: azotu (—147°C) i tlenu (—119°C). Dodatkowe oziębienie uzyskuje się przez zmniejszenie ciśnienia nad skroplonym etylenem. Pod ciśnieniem około 100 mm Hg temperatura wrzenia etylenu wynosi około —150°C. Oziębienie powietrza, azotu lub tlenu do tej temperatury, przy równoczesnym zastosowaniu zwiększonego ciśnienia, wystarcza do skroplenia wymienionych gazów.

Skroplone powietrze poddane ciśnieniu 1 atm oziębia się wskutek gwałtownego parowania do —191°C. Zmniejszając ciśnienie w większym stopniu można uzyskać dalsze obniżenie temperatury wrzenia. Wróblewski i Olszewski doszli w ten sposób do temperatury —220°C, niewystarczającej zresztą jeszcze do skroplenia wodoru i helu.



Rys. 17.15

Obecnie skraplanie powietrza odbywa się *metodą Hampsona* i *Lindego*. Metoda ta wykorzystuje oziębianie się gazów podczas gwałtownego rozprężania. Schemat przyrządu Hampsona przedstawia rys. 17.15.

Powietrze sprężone w kilkustopniowej sprężarce do ciśnienia około 200 atm przepływa przez wąską miedzianą rurkę zwiniętą w wężownicę. Dolny wylot rurki może być dowolnie zmniejszany przez wprowadzenie do otworu stożkowatego zakończenia śruby *B*. Jeśli szerokość otworu jest niewielka, w rurce miedzianej utrzymuje się ciśnienie bliskie 200 atm. Na zewnątrz rurki panuje ciśnienie 1 atm. Powietrze wydostając się z wylotu gwałtownie się rozpręża i równocześnie silnie się oziębia. Masy powietrza o niskiej temperaturze wędrują między zwojami rurki miedzianej do wylotu *D*, oziębiając przy tym zawartość rurki. Dzięki temu dalsze porcje powietrza dochodzą do wylotu *A* już ochłodzone, a w czasie rozprężania oziębiają się jeszcze bardziej.

Ten proces odbywa się stale doprowadzając do coraz większego obniżenia temperatury powietrza. Wreszcie pewna porcja powietrza dochodząca do wylotu *A* osiąga tak niską temperaturę podczas rozprężania, że ulega skropleniu. Skroplone powietrze zbiera się w dolnej części zbiornika *Z*, skąd przez otwarcie kranu *K* jest wylewane do naczynia Dewara *N*.

Obie metody łącznie zastosowano do skraplania dwóch gazów o najniższych temperaturach krytycznych, a mianowicie wodoru i helu. Gazy te oziębiają się przy rozprężaniu dopiero wtedy, gdy są dostatecznie oziębione: wodór poniżej —80,5°C, hel poniżej —248,5°C. Trzeba więc przed wprowadzeniem do przyrządu Hampsona i Lindego obniżyć temperaturę wodoru np. za pomocą ciekłego powietrza, a temperaturę helu za pomocą ciekłego wodoru.

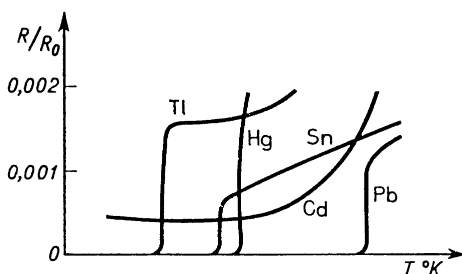
W ten sposób Dewar po raz pierwszy skroplił wodór w większej ilości w r. 1898, a Kammerlingh Onnes skroplił hel w r. 1908. Ciekły hel pod normalnym ciśnieniem wrze w temperaturze 4,3K. Zmniejszając ciśnienie nad helem, uzyskano temperaturę

0,7K. Osiągnięcie niższych jeszcze temperatur stało się możliwe dzięki wykorzystaniu właściwości ciał paramagnetycznych. Najniższa temperatura uzyskana w laboratorium wynosi 0,0044K.

Ostatnio opracowano nową metodę skraplania gazów. Zasada jej polega na tym, że gaz silnie sprężony wprowadzony zostaje do zbiornika zamkniętego tłokiem. Podczas przesuwania tłoka gaz wykonuje pracę pokonując siły zewnętrzne. Praca ta powstaje kosztem energii wewnętrznej gazu. Towarzyszy temu obniżenie jego temperatury doprowadzające ostatecznie do skroplenia (metoda opracowana przez fizyka radzieckiego Kapicę).

17.11. Właściwości ciał w niskich temperaturach

W niskich temperaturach liczne właściwości ciał ulegają poważnym zmianom. W temperaturze ciekłego powietrza radykalnie zmieniają się właściwości sprężyste niektórych ciał. Tak np. kauczuk staje się kruchy jak szkło, rtęć występuje w stanie stałym i jest kowalna, maleją wartości ciepła właściwego, oporu elektrycznego przewodników itp.



Rys. 17.16

Interesującą właściwość wykazują niektóre metale w temperaturach bliskich zera bezwzględnego. Opór właściwy większości metali, jak wykazały badania Kammerlingha Onnesa, maleje stopniowo w miarę zmniejszania się temperatury, ale tylko do pewnej temperatury granicznej. Począwszy od tej temperatury wartość oporu właściwego pozostaje stała. Niektóre metale wykazują jednak gwałtowny spadek oporu praktycznie do zera w temperaturach bliskich zera bezwzględnego. Mówimy, według określenia Kammerlingha Onnesa, że metal w tej temperaturze staje się ***nadprzewodnikiem***. Nadprzewodnictwo zostało odkryte w przypadku rtęci, ołowiu, cynku, kadmu i wielu innych metali. Natomiast nie odkryto nadprzewodnictwa między innymi: miedzi, srebra, złota, żelaza, sodu, potasu. Rys. 17.16 przedstawia zależność oporu różnych metali (R/R_0 — stosunek oporów danego przewodnika w temperaturze badanej i 0°C) od temperatury w temperaturach bliskich zera bezwzględnego.

Badania uczonego radzieckiego Kapicy dały interesujące wyniki dotyczące właściwości helu w temperaturach bliskich 0K. Hel skrapla się pod normalnym ciśnieniem w temperaturze 4,8K. Poddany następnie zmniejszonemu ciśnieniu wrze i oziębia się. Ten

proces trwa do chwili osiągnięcia temperatury 2,19K. W tej temperaturze wrzenie ustaje. Hel pozostaje w stanie ciekłym, ale przekształca się na odmianę zwaną hel II. W czasie tej przemiany ulegają gwałtownej zmianie gęstość, ciepło właściwe, stała dielektryczna itd. Najbardziej zdumiewającą właściwością helu II jest jego zdolność szybkiego przenoszenia ciepła. Celem wytłumaczenia tej właściwości Kapica zakłada, że hel II jest pozbawiony lepkości. Założenie to potwierdzają doświadczenia z przepływem helu II przez rurki włoskowate. Wobec braku lepkości w helu II powstają prądy konwekcyjne pośredniczące w szybkim przekazywaniu ciepła. Tę właściwość helu II Kapica zaproponował nazywać nadciekłością.

17.12. Równanie stanu gazów rzeczywistych

Po omówieniu właściwości par i zestawieniu badań doświadczalnych z tej dziedziny wrócimy raz jeszcze do opisu właściwości gazów rzeczywistych.

Wiemy, że równanie Clapeyrona stosuje się do gazów rzeczywistych tylko z pewnym przybliżeniem. Wielu uczonych usiłowało dojść do równania opisującego dokładniej zależność między wielkościami p, V i T w odniesieniu do gazów rzeczywistych. Jedna z takich bardziej udanych prób wiąże się z nazwiskiem van der Waalsa.

Celem dokładniejszego scharakteryzowania właściwości gazu rzeczywistego van der Waals uwzględnił zaniedbane w przypadku gazu doskonałego:

1) siły międzycząsteczkowe gazu,

2) objętość własną cząsteczek gazu.

Uwzględnienie sił spójności między cząsteczkami gazu sprowadza się do wyznaczenia poprawki na ciśnienie. Zwróćmy uwagę na dowolną cząsteczkę gazu x, dobiegającą do ściany naczynia. Wobec braku cząsteczek gazu od strony ściany wypadkowe działanie na cząsteczkę x jest skierowane w głąb naczynia. Wpływa ono hamująco na ruch cząsteczki zmniejszając jej prędkość, pęd, a zatem i ciśnienie na ściankę. Gdybyśmy zmierzyli ciśnienie działające na ścianki, otrzymalibyśmy z pomiaru wartość za małą w porównaniu z ciśnieniem panującym wewnątrz naczynia. Aby otrzymać wartość ciśnienia wewnętrznego, trzeba do ciśnienia zmierzonego dodać pewną poprawkę. Według van der Waalsa poprawka ta ma postać a/V_m^2, gdzie a jest stałą charakterystyczną dla danego gazu rzeczywistego.

W równaniu $pV_m = RT$ wielkość V_m oznacza objętość naczynia zawierającego 1 mol gazu. Uwzględniając objętości własne cząsteczek musimy przyjąć, że objętość, w której mogą się poruszać cząsteczki gazu, jest zmniejszona o pewną wartość zależną od objętości własnej cząsteczek. Z rozważań van der Waalsa wynika, że ta druga poprawka b (stała dla danego gazu) powinna być równa czterokrotnej objętości własnej cząsteczek wchodzących w skład jednego mola.

Ostatecznie równanie van der Waalsa przyjmuje postać:

$$\left(p+\frac{a}{V_m^2}\right)(V_m-b) = RT. \tag{17.4}$$

Poprawka b, z natury swej niewielka, staje się bardzo mała w porównaniu z V_m

zarówno wtedy, gdy gaz jest ogrzany do wysokiej temperatury, jak i w przypadku, gdy mamy do czynienia z niewielkimi ciśnieniami. Ze wzrostem V_m maleje też wyrażenie a/V_m^2. W tych warunkach

$$p \gg \frac{a}{V_m^2} \quad \text{i} \quad V_m \gg b,$$

możemy więc obie poprawki zaniedbać. Innymi słowy, w wysokiej temperaturze i w odniesieniu do niewielkich ciśnień równanie van der Waalsa redukuje się do wyrażenia $pV_m = RT$, tzn. do równania Clapeyrona.

Do dalszej analizy równania van der Waalsa lepiej się nadaje inna jego postać. Sprowadzając do wspólnego mianownika i otwierając nawiasy otrzymamy

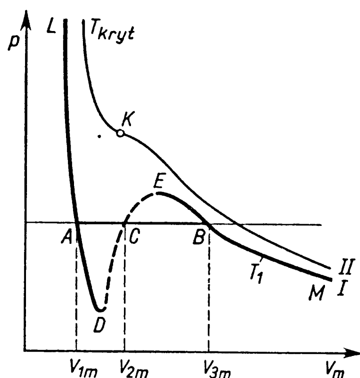
$$pV_m^3 + aV_m - bpV_m^2 - ab - RTV_m^2 = 0,$$

czyli

$$pV_m^3 - (bp + RT)V_m^2 + aV_m - ab = 0$$

lub

$$V_m^3 - \left(b + \frac{RT}{p}\right)V_m^2 + \frac{a}{p}V_m - \frac{ab}{p} = 0. \tag{17.5}$$



Rys. 17.17

Jest to równanie trzeciego stopnia względem objętości molowej V_m. Dla danych p i T równanie to w zasadzie powinno mieć trzy pierwiastki. Aby zbadać ich znaczenie fizyczne, wykreślamy krzywe zależności p i V_m dla stałych temperatur (rys. 17.17). Okazuje się, że w temperaturach niższych od krytycznych krzywa ma przebieg izotermy I, a w temperaturze krytycznej — przebieg izotermy II. Dla izotermy I charakterystyczne jest to, że w pewnym przedziale wartości p linia prosta odpowiadająca stałemu ciśnieniu, a więc równoległa do osi odciętych, przecina krzywą w trzech punktach wyznaczających trzy rzeczywiste pierwiastki równania. Niech np. prosta AB odpowiada wartości prężności pary nasyconej danej cieczy w temperaturze T_1. Wtedy rzeczywistymi pierwiastkami równania van der Waalsa będą trzy różne objętości molowe: V_{1m}, V_{2m}, V_{3m}.

Chcemy określić znaczenie fizyczne tych trzech objętości molowych. W tym celu raz jeszcze wykorzystujemy wykresy Andrewsa korzystając z danych doświadczalnych. Porównanie krzywej teoretycznej van der Waalsa (rys. 17.17) z krzywą doświadczalną Andrewsa (rys. 17.13) doprowadza do następujących wniosków. Przebieg krzywej van der Waalsa na odcinku *LA* i przebieg krzywej Andrewsa na odcinku *CA* cieczy jest jednakowy. Punkt *A* na wykresie Andrewsa ma odciętą, wyrażającą maksymalną objętość danej masy substancji badanej występującej jeszcze w całości w stanie ciekłym. Przez analogię zatem powiemy, że V_{1m} na wykresie van der Waalsa przedstawia maksymalną objętość molową cieczy w temperaturze T_1, dla której została wykreślona izoterma *I*.

Porównanie odcinka *BM* z odcinkiem pary nienasyconej na wykresie Andrewsa pozwala na ustalenie znaczenia V_{3m}. Jest to objętość molowa pary nasyconej suchej w temperaturze T_1. Przebieg krzywej doświadczalnej i teoretycznej między punktami *A* i *B* jest różny. Zastanówmy się, czy poszczególne części krzywej *ADCEB* mogą odpowiadać realnym stanom. Znamy już pewne stany „przekroczenia równowagi". Wiemy, że podczas ostrożnego oziębiania czystej substancji ciekłej można ją oziębić poniżej temperatury krzepnięcia bez wywołania zestalenia. Otrzymuje się wtedy ciecz przechłodzoną. Można też zachowując odpowiednią ostrożność ogrzać ciecz powyżej temperatury wrzenia, bez rozpoczęcia procesu wrzenia. Pewnym stanom przekroczenia mogą właśnie odpowiadać punkty krzywej van der Waalsa na odcinku *AD* oraz na odcinku *EB*. Odcinek *AD* odpowiadałby cieczy przegrzanej, odcinek *EB* — parze przesyconej. Para przesycona mogłaby powstać, gdybyśmy zmniejszając nieskończenie powoli objętość pary nienasyconej przekroczyli stan pary nasyconej suchej w punkcie *B* i w dalszym ciągu zmniejszając objętość nie wywołali procesu skraplania.

Tak więc odcinki *AD* i *EB* mogłyby odpowiadać pewnym wyjątkowym, nietrwałym, ale jednak realnym stanom. Inaczej się sprawa przedstawia z odcinkiem *DE* krzywej. Punkty tej części krzywej odpowiadają takim przemianom, w których ze wzrostem objętości rośnie ciśnienie. Takie przemiany są niemożliwe. Ta część krzywej nie ma żadnego realnego odpowiednika, a zatem punkt *C* i odpowiadająca mu objętość molowa V_{2m} nie mają żadnego realnego znaczenia fizycznego.

Przebieg krzywej van der Waalsa w temperaturze krytycznej jest analogiczny do przebiegu izotermy krytycznej z wykresu Andrewsa. Odcięta punktu *K* przegięcia krzywej odpowiada potrójnemu pierwiastkowi równania

$$V_{m\,\text{kryt}} = V_{1m} = V_{2m} = V_{3m}.$$

Uwzględniając pierwiastek potrójny równania van der Waalsa, czyli równania trzeciego stopnia, równanie to można napisać w postaci

$$(V_m - V_{m\,\text{kryt}})^3 = 0,$$

co po rozwinięciu daje

$$V_m^3 - 3V_{m\,\text{kryt}}V_m^2 + 3V_{m\,\text{kryt}}^2 V_m - V_{m\,\text{kryt}}^3 = 0. \tag{17.6}$$

Wprowadzając do ogólnego równania van der Waalsa parametry T_{kryt} i p_{kryt}, dotyczące stanu krytycznego, otrzymamy

$$V_m^3 - \left(b + \frac{RT_{\text{kryt}}}{p_{\text{kryt}}}\right)V_m^2 + \frac{a}{p_{\text{kryt}}}V_m - \frac{ab}{p_{\text{kryt}}} = 0. \tag{17.7}$$

Z porównania współczynników przy tych samych potęgach V_m w równaniach (17.6) i (17.7) wynikają następujące związki:

$$3V_{m\,\text{kryt}} = b + \frac{RT_{\text{kryt}}}{p_{\text{kryt}}}, \tag{17.8}$$

$$3V_{m\,\text{kryt}}^2 = \frac{a}{p_{\text{kryt}}}, \tag{17.9}$$

$$V_{m\,\text{kryt}}^3 = \frac{ab}{p_{\text{kryt}}}. \tag{17.10}$$

Z tego układu trzech równań znajdujemy następujące zależności między a, b, R i parametrami stanu krytycznego — objętością, ciśnieniem i temperaturą:

$$V_{m\,\text{kryt}} = 3b,$$

$$p_{\text{kryt}} = \frac{a}{27b^2},$$

$$T_{\text{kryt}} = \frac{8a}{27bR}.$$

Można te związki stosować do obliczania poprawek van der Waalsa w oparciu o wyznaczone doświadczalnie parametry stanu krytycznego lub odwrotnie — znając a, b, R można obliczać parametry krytyczne.

Krzywe van der Waalsa wykreślone dla temperatur wyższych od krytycznych będą miały tylko jeden punkt przecięcia z prostymi odpowiadającymi $p = \text{const}$. W takich warunkach istnieje tylko jeden pierwiastek rzeczywisty równania i dwa zespolone sprzężone.

Interesującą postacią równania van der Waalsa jest tzw. ***równanie zredukowane***. Zostało ono wyprowadzone przy założeniu, że stanami odpowiadającymi sobie w przypadku różnych ciał są ich stany krytyczne. Wobec tego, wyrażając wartości a, b i R za pośrednictwem charakterystycznych dla danego ciała parametrów krytycznych i podstawiając otrzymane wyrażenia do równania van der Waalsa, otrzymamy nową, ogólniejszą jego postać. A mianowicie łatwo znajdziemy, że

$$b = \frac{V_{\text{kryt}}}{3}, \qquad a = 3V_{\text{kryt}}^2 p_{\text{kryt}}, \qquad R = \frac{8V_{\text{kryt}}p_{\text{kryt}}}{3T_{\text{kryt}}}.$$

Podstawiając te wartości do równania (17.4) otrzymamy

$$\left(p + \frac{3V_{\text{kryt}}^2 p_{\text{kryt}}}{V_m^2}\right)\left(V_m - \frac{V_{\text{kryt}}}{3}\right) = \frac{8V_{\text{kryt}}p_{\text{kryt}}}{3T_{\text{kryt}}}T.$$

Podzielenie obu stron przez iloczyn $p_{\text{kryt}}V_{\text{kryt}}$ daje

$$\left(\frac{p}{p_{\text{kryt}}} + \frac{3}{(V_m/V_{\text{kryt}})^2}\right)\left(\frac{V_m}{V_{\text{kryt}}} - \frac{1}{3}\right) = \frac{8T}{3T_{\text{kryt}}}.$$

Wprowadzając wielkości

$$p/p_{\text{kryt}} = \pi, \quad V_m/V_{\text{kryt}} = \sigma, \quad T/T_{\text{kryt}} = \vartheta$$

porównujemy aktualne ciśnienie, objętość i temperaturę badanego ciała z charakterystycznymi dla niego parametrami krytycznymi (innymi słowy, aktualne parametry wyrażamy w sposób względny, przyjmując za jednostki odpowiednie wartości parametrów krytycznych). Po podstawieniu wyżej wprowadzonych symboli zredukowane równanie van der Waalsa przyjmuje postać:

$$\left(\pi+\frac{3}{\sigma^2}\right)(\sigma-\tfrac{1}{3}) = \tfrac{8}{3}\vartheta \tag{17.11}$$

i nie zawiera już żadnych stałych zależnych od rodzaju ciał. Z równymi wartościami π, σ i ϑ różnych ciał związane są ich tzw. *stany odpowiednie*. W związku z tym ostatnio otrzymane równanie nazywane jest czasem *równaniem stanów odpowiednich*.

Powyższe równanie powinno stosować się do gazów i do cieczy, jeśli słuszna jest hipoteza ciągłości przejścia między tymi fazami. Doświadczenia potwierdziły z pewną dokładnością te przypuszczenia w odniesieniu do dość dużej liczby cieczy. W tych przypadkach, gdy równanie wyraźnie się nie stosowało, stwierdzono na podstawie innych badań, że w cieczach tworzyły się układy zasocjowane, czyli że skład cząsteczek cieczy był inny niż skład cząsteczek gazu.

17.13. Wykresy równowagi fazowej. Prawo Clausiusa-Clapeyrona. Punkt potrójny

Wróćmy raz jeszcze do podstawowego tematu niniejszego rozdziału — do przemian fazowych.

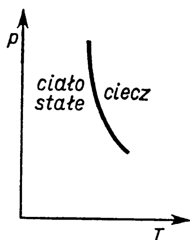
Mówiąc o temperaturze topnienia podkreślaliśmy, że temperatura topnienia danej substancji jest to temperatura, w której obie fazy, stała i ciekła, pozostają w równowadze ze sobą; gdy nie działają na taki układ siły zewnętrzne, to ani ciecz nie krzepnie, ani ciało stałe się nie topi. Ale — jak pamiętamy — ta temperatura jest funkcją ciśnienia. Można zatem zrobić tzw. *wykres fazowy* we współrzędnych T i p, otrzymując na nim krzywą równowagi układu ciecz-ciało stałe (krzywa krzepnięcia) (rys. 17.18), rozdzielającą płaszczyznę wykresu na obszar ciała stałego i cieczy. W analogiczny sposób można wykreślić krzywą równowagi układu ciecz-gaz (rys. 17.19). Każdy punkt otrzymanej krzywej (zwanej krzywą parowania) odpowiada warunkom, w których ciecz i para danej substancji mogą istnieć w równowadze ze sobą: ciecz nie paruje, para się nie skrapla (w ujęciu makroskopowym). Znowu nasza krzywa rozdziela płaszczyznę wykresu na obszar cieczy i obszar gazu. Obszar cieczy przy każdym określonym ciśnieniu leży po stronie niższych temperatur (a więc po lewej stronie krzywej parowania), a obszar pary — po stronie wyższych temperatur (a więc po stronie prawej krzywej parowania).

W analogiczny sposób można by wykreślić krzywą równowagi dla fazy stałej i gazowej (krzywa sublimacji).

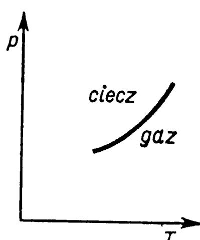
Nachylenia krzywych równowagi fazowej we współrzędnych $p = p(T)$ określa prawo Clausiusa–Clapeyrona, według którego

$$\frac{dp}{dT} = \frac{q}{T(V_2 - V_1)}, \tag{17.12}$$

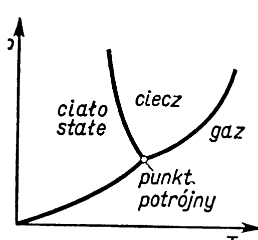
gdzie q oznacza ciepło danej przemiany fazowej, a więc ciepło topnienia, ciepło parowania, ciepło sublimacji, $(V_2 - V_1)$ — różnicę objętości tej samej masy, do której odniesione jest ciepło przemiany, w obu rozważanych fazach w temperaturze T.



Rys. 17.18



Rys. 17.19



Rys. 17.20

Ze wzoru Clausiusa–Clapeyrona można bezpośrednio wyciągać wnioski dotyczące stromości przebiegu krzywych równowagi w przypadku np. topnienia i parowania. W przypadku parowania różnica objętości $V_2 - V_1$ jest znacznie większa niż w przypadku topnienia, a zatem nachylenie krzywej parowania (określone przez dp/dT) jest znacznie mniejsze. Innymi słowy, krzywa parowania przebiega mniej stromo: mniejsza zmiana ciśnienia potrzebna jest do wywołania określonej zmiany temperatury wrzenia, niż to ma miejsce w przypadku zmiany temperatury topnienia. Ilustracją liczbową podanych zależności będą następujące dane: dla wywołania zmiany o 1 K temperatury wrzenia wody i topnienia lodu potrzebne są odpowiednio zmiany ciśnienia równe 27 mm Hg i 130 atm (około 3600 N/m^2 i około $13 \cdot 10^6$ N/m^2).

Podane wyżej rys. 17.18 i 17.19 dotyczyły równowagi fazowej dwóch faz. Nasuwa się pytanie, w jakich warunkach zachodzić może równowaga trzech faz? Równowaga trzech faz możliwa jest jedynie w warunkach odpowiadających wspólnemu punktowi omówionych trzech wykresów fazowych (z wykresem sublimacji włącznie) (rys. 17.20). Punkt ten nosi nazwę *punktu potrójnego*. Ściśle określone wartości temperatury i ciśnienia wyznaczają warunki równowagi trzech faz danej substancji. Dla wody punkt potrójny określony jest wartościami $p = 4{,}62$ mm Hg i $t = 0{,}01°C$, czyli 273,16K. Wykorzystaliśmy już te dane przy określaniu jednostki temperatury w układzie SI (por. § 3 „Wiadomości wstępnych”).

Pytania i zadania

1. Jakie procesy zaliczamy do przemian fazowych? Podać ich charakterystykę i przykłady.
2. Scharakteryzować proces topnienia ciał krystalicznych.
3. Zdefiniować temperaturę topnienia. Wyjaśnić różną zależność temperatury topnienia od ciśnienia dla dwóch grup ciał.

4. Jaka jest definicja ciepła topnienia? W jakich jednostkach ono się mierzy? Jak brzmiałaby definicja molowego ciepła topnienia?

5. Scharakteryzować zjawisko parowania. Rozróżnić zjawisko parowania i wrzenia. Jaka jest zależność temperatury wrzenia od ciśnienia? Czemu się równa ciśnienie pary nasyconej w temperaturze wrzenia?

6. Podać definicję ciepła parowania. Jak można określić molowe ciepło parowania?

7. Którą z par (parę nasyconą czy nienasyconą) charakteryzuje ściśle określona prężność w danej temperaturze? W jakich granicach zmienia się prężność pary drugiego rodzaju w danej temperaturze?

8. W jaki sposób parę nienasyconą można przekształcić w parę nasyconą?

9. Która z par ma właściwości zupełnie odmienne od właściwości gazów rzeczywistych? Do czego sprowadzają się różnice?

10. Dlaczego prężność pary nasyconej ze wzrostem temperatury szybciej rośnie niż prężność pary nienasyconej?

11. Omówić badania doświadczalne, np. dwutlenku węgla prowadzące do graficznego przedstawienia $p = p(V)$ przy $T =$ const dla tego związku w postaci cieczy, pary nasyconej i nienasyconej.

12. O czym świadczy występujący na izotermie (rys. 17.13) w temperaturze niższej od krytycznej: a) prostoliniowy odcinek AB, b) hiperboliczny odcinek odpowiadający parze nienasyconej? O czym świadczy zanik odcinka AB w temperaturze krytycznej?

13. Co jest warunkiem koniecznym do skroplenia gazów?

14. Jakie właściwości wykazują ciała w temperaturach dużo niższych od 0°C, w temperaturze bliskiej 0K?

15. Czym się różni gaz doskonały od gazu rzeczywistego?

16. Jakie poprawki wprowadził van der Waals do równania Clapeyrona?

17. Czy wszystkie trzy rzeczywiste pierwiastki równania van der Waalsa mają sens fizyczny? Omówić wykres równania, sens fizyczny poszczególnych części krzywych oraz znaczenie pierwiastków V_{1m}, V_{2m}, V_{3m}.

18. Na czym polega koncepcja stanów podobnych?

19. Co przedstawiają i jaki mają przebieg krzywe równowagi fazowej układów dwuskładnikowych?

20. Czy równowaga trzech faz możliwa jest przy kilku różnych zestawach wartości ciśnienia i temperatury? Co oznacza punkt potrójny?

21. Jaki jest związek punktu potrójnego wody z określeniem kelwina?

22. Czy znając przebieg krzywej równowagi oraz objętość 1 kg danego ciała w fazie ciekłej i stałej w pewnej temperaturze T można obliczyć ciepło topnienia w tej temperaturze? Jakie prawo jest tutaj przydatne?

23. Do kalorymetru o masie m_k i cieple właściwym c_k zawierającego m_w wody o cieple właściwym c_w i temperaturze t_0 wlano m_x stopionego metalu o temperaturze t_x wyższej od temperatury topnienia θ i od temperatury t_0. W kalorymetrze ustaliła się pewna temperatura końcowa t_k, przy czym $t_0 < t_k < \theta < t_x$. Wypisać równanie bilansu cieplnego i wyznaczyć ciepło topnienia metalu, przyjmując, że ciepło właściwe metalu w stanie ciekłym i stałym wynosi odpowiednio c_x' i c_x''.

24. Do masy $m = 500$ g wody o temperaturze $t_0 = 50°C$ wrzucono $m_1 = 50$ g lodu o temperaturze $t_1 = -10°C$. Obliczyć temperaturę końcową. Potrzebne dane liczbowe można znaleźć w tablicach.

Odp. 38°C.

25. Ile ciepła potrzeba do ogrzania, stopienia i wyparowania w temperaturze 100°C bryły lodu o temperaturze początkowej −10°C i masie 1 kg.

Odp. $3{,}03 \cdot 10^6$ J.

26. Wiedząc, że ciepło parowania wody pod ciśnieniem normalnym równa się $226 \cdot 10^4$ J/kg i przyjmując, że podczas wrzenia w warunkach normalnych objętość 1 kg wody zmienia się z 10^3 cm³ na 1671×10^3 cm³, obliczyć, jaka część doprowadzonego ciepła zużywa się na wykonanie pracy zewnętrznej związanej ze wzrostem objętości, a jaka część na pracę wewnętrzną związaną z pokonaniem wzajemnego przyciągania między cząsteczkami wody.

Odp. 0,076 i 0,924.

27. Stała van der Waalsa dla tlenu $b = 32$ cm³/mol. Zakładając, że cząsteczki tlenu mają kształt ku-

listy i pamiętając, że stała b odpowiada czterokrotnej objętości własnej cząsteczek, wyznaczyć promień cząsteczki tlenu.

28. Wyznaczyć stałe van der Waalsa dla CO_2, wiedząc, że jego ciśnienie krytyczne równa się 73 at, a temperatura krytyczna 31,1°C.

29. Obliczyć ciśnienie, jakie wywiera masa 1,1 kg CO_2 zamkniętego w zbiorniku o pojemności 20 l w temperaturze 13°C. Stałe van der Waalsa dla dwutlenku węgla: $a = 0{,}37$ Nm^4/mol^2, $b = 43 \times 10^{-6}$ m^3/mol. Porównać to ciśnienie z ciśnieniem obliczonym przy założeniu stosowalności prawa Clapeyrona.

Odp. 25 at, 29 at.

30. Obliczyć ciepło topnienia ołowiu pod normalnym ciśnieniem ($\theta = 327°C$), jeżeli wiadomo, że przy wzroście ciśnienia o 1 atm temperatura topnienia wzrasta o 0,008°C i że towarzyszący topnieniu wzrost objętości wynosi 3,4%. Gęstość ołowiu 11,3 g/cm^3.

Odp. $22{,}86 \cdot 10^3$ J/kg.

31. W stalowej kuli o pojemności 0,53 m^3 znajduje się 1 kmol dwutlenku węgla pod ciśnieniem 50 atm. Jaka jest temperatura tego gazu. Stałe van der Waalsa jak w zadaniu 29.

Odp. 100°C.

OZDZIAŁ 18

Termodynamika

18.1. Uwagi wstępne

Termodynamika jest to dział fizyki, w którym zjawiska cieplne są rozpatrywane z punktu widzenia energetyki, ze szczególnym uwzględnieniem możliwości i warunków przemiany ciepła na pracę i odwrotnie.

W wieku XVIII i w pierwszej połowie wieku XIX ciepło było uważane za rodzaj nieważkiej materii, zwanej cieplikiem. W oparciu o teorię cieplikową można było tłumaczyć zjawiska wymiany ciepła między ciałami, jako przekazywanie cieplika od jednego ciała do drugiego związane ze spadkiem temperatury pierwszego ciała, a wzrostem drugiego. Wystąpiły jednak trudności przy tłumaczeniu takich zjawisk, podczas których wywiązuje się ciepło kosztem pracy. Do takich zjawisk należało np. silne ogrzewanie się ciał przy wierceniu otworów luf armatnich (zaobserwowane przez Rumforda), jak również zjawisko topienia się dwóch kawałków lodu w wyniku ich wzajemnego pocierania (zaobserwowane przez Davy'ego).

Te i tym podobne zjawiska dostarczyły podstaw do zwalczania teorii cieplikowej. Na tym tle powstaje pogląd, że ciepło jest rodzajem energii, może powstawać kosztem pracy i może być zamieniane na pracę.

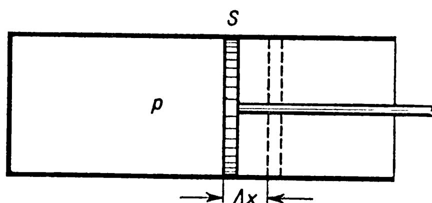
Jak już wspominaliśmy w rozdz. 15, przyjęcie układu SI jako obowiązującego układu jednostek zdecydowało o wycofaniu z użycia specjalnej jednostki ciepła — kalorii, pozostawiając jako obowiązującą jednostkę pracy i energii (a więc i ilości ciepła) 1 dżul i jego pochodne. Straciły też znaczenie takie wielkości, jak mechaniczny równoważnik ciepła równy 4,186 J/cal lub 426,8 kGm/cal oraz cieplny równoważnik pracy równy 0,239 cal/J. Przypominamy je jednak z tego względu, że w wielu podręcznikach, pracach naukowych, tabelach wielkości fizykochemicznych występują jeszcze kalorie. Znajomość równoważników umożliwia przeliczanie kalorii na dżule i odwrotnie.

18.2. Wprowadzenie podstawowych pojęć z dziedziny termodynamiki

Jednym z podstawowych terminów, którym stale będziemy się posługiwali w termodynamice, jest pojęcie *układu termodynamicznego* (zwanego też ciałem roboczym,

czynnikiem termodynamicznym lub krótko — czynnikiem). Przykładem układu termodynamicznego może być np. gaz w zbiorniku zamkniętym tłokiem.

W układzie termodynamicznym mogą zachodzić różnego rodzaju przemiany termodynamiczne. Jeśli przemiana jest takiego rodzaju, że stan początkowy i końcowy układu są jednakowe, mówimy o *przemianie kołowej* (albo *przemianie zamkniętej*, *cyklu*, *obiegu*). Jeśli stan początkowy i końcowy układu są różne, mówimy o *przemianie otwartej*.



Rys. 18.1

Formy energetycznego oddziaływania badanego układu termodynamicznego z jego otoczeniem mogą być różne: sprowadzają się one do wymiany ciepła i przekazywania pracy. Ujmujemy to krótko mówiąc o dostarczaniu układowi ciepła lub pracy, ewentualnie o oddawaniu przez układ ciepła i wykonywaniu przez układ pracy. Tak np. przy rozprężaniu się gazu zamkniętego w cylindrze o przekroju S (rys. 18.1) pod ciśnieniem p zostanie przez układ wykonana praca ΔW:

$$\Delta W = pS\Delta x = p\Delta V.$$

Przy wypisywaniu równań używanych w termodynamice stosuje się pewną umowę co do znaków przypisywanych wartościom ΔW i ΔQ. Pracę uważamy za ujemną, jeżeli jest ona wykonywana przez siły zewnętrzne na układzie (praca dostarczana układowi). Odwrotnie, ΔW jest dodatnie, gdy układ wykonuje pracę dzięki działaniu sił wewnętrznych. Tak więc rozważana przez nas praca wykonana podczas rozprężania się gazu w cylindrze (rys. 18.1) jest pracą dodatnią.

Ciepło dostarczane do układu będziemy uważali za dodatnie, natomiast ciepło oddawane przez układ — za ujemne. Jeśli zatem zajdzie pewna przemiana termodynamiczna, przy której

$$\Delta Q = \Delta W$$

(przy czym obie wartości ΔQ i ΔW są dodatnie), to zgodnie z wprowadzonymi umowami równanie to odczytamy w ten sposób, że kosztem dostarczonego ciepła układ wykonuje pracę (*obieg* tzw. *prosty*). Jeśli natomiast obie wartości ΔQ i ΔW są ujemne, to przytoczone równanie odczytamy tak, że kosztem dostarczonej pracy układ oddaje ciepło (*obieg* tzw. *odwrotny*).

Bardzo ważnym pojęciem w termodynamice jest energia wewnętrzna układu (omówieniu jej poświęcimy § 18.4). W tej chwili podkreślamy, że przemiany kołowe, w któ-

rych stan początkowy i końcowy układu są identyczne, zachodzą oczywiście bez zmian energii wewnętrznej. Do takich właśnie przemian odnoszą się równania typu

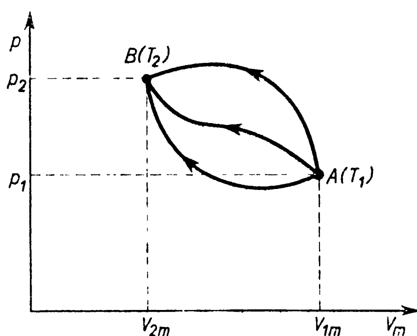
$$\Delta Q = \Delta W.$$

Zwrącać jednak należy uwagę na poprawną ocenę tego, czy przemiana jest kołowa. Rozważmy np. przypadek, gdy przesuwając tłok zmniejszamy objętość pary nasyconej, czyli powodujemy jej częściowe skroplenie. Podczas tej przemiany dostarczyliśmy układowi pewną pracę ΔW. Skraplająca się para oddała pewną ilość ciepła ΔQ. Jednakże rozważany układ uległ zmianie: para nasycona zmieniła się częściowo na ciecz. Dostarczana układowi praca ΔW nie jest równa oddanej ilości ciepła ΔQ.

18.3. Graficzne przedstawienie przemian termodynamicznych zachodzących w układzie

Doprowadzanie ciepła, jak również wykonywanie pracy przez układ, wywołuje zmiany parametrów charakteryzujących stan danego układu. W termodynamice stan układu można określać np. za pomocą trzech parametrów p, V i T. Nie są to jednak wielkości od siebie niezależne. Dla różnych ciał znamy równania wiążące je ze sobą: np. równanie Clapeyrona dla gazu doskonałego, równanie van der Waalsa dla gazu rzeczywistego itp. Jeżeli znamy dwa parametry oraz wzór wyrażający zależność między nimi, to wartość trzeciego wynika bezpośrednio z równania wiążącego p, V i T. Tak więc w układzie współrzędnych (V_m, p) punkt A (rys. 18.2) o współrzędnych (V_{1m}, p_1) określa jednoznacznie stan jednego mola gazu doskonałego, zajmującego objętość V_{1m} pod ciśnieniem p_1 w temperaturze T_1, gdyż temperatura T_1 wynika z równania

$$p_1 V_{1m} = RT_1.$$



Rys. 18.2

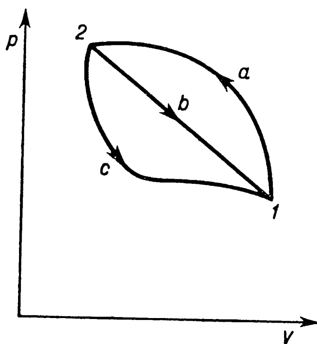
Punkt B wykresu odpowiada innemu stanowi tego mola gazu, a mianowicie wartościom V_{2m} i p_2, odczytanym bezpośrednio z wykresu oraz wartości T_2, obliczonej z równania Clapeyrona.

Przejście od stanu A do stanu B może się odbywać w sposób ciągły przez szereg stanów przejściowych. Odpowiada to na wykresie różnym liniom łączącym A z B.

Linie te mogą mieć dowolny kształt, zależny od rodzaju przemiany — przejścia od A do B. Takie wykreślne przedstawienie przemian jest w termodynamice szeroko stosowane.

18.4. Energia wewnętrzna

Aby przejść do ogólnego wysłowienia pierwszej zasady termodynamiki, musimy skorzystać z pojęcia energii wewnętrznej wprowadzonego do nauki przez Kelvina. Na ogólny zasób energii dowolnego spoczywającego układu składają się: energia potencjalna, cieplna, chemiczna, magnetyczna, jądrowa itp. Te wszystkie rodzaje energii obejmujemy wspólną nazwą energii wewnętrznej. W termodynamice nie interesują nas formy występowania energii wewnętrznej. Jest obojętne, w jakim stopniu energia ta wiąże się z energią kinetyczną ruchu molekularnego lub z energią potencjalną wzajemnego oddziaływania cząsteczek, z energią wiązań chemicznych, z energią atomową itd. Interesują nas tylko ilościowe zmiany energii wewnętrznej, towarzyszące przejściu układu z jednego stanu do drugiego. Trzeba sobie zdawać sprawę, że ogólnego zasobu energii wewnętrznej we wszystkich postaciach nie potrafimy wyznaczyć, podobnie jak nie umiemy podać ogólnego zasobu energii potencjalnej ciała wzniesionego na wysokość h nad poziom ziemi. Potrafimy tylko ustalać różnice wartości energii potencjalnej, odpowiadające dwóm różnym poziomom.



Rys. 18.3

Różnicę energii wewnętrznej ΔU, odpowiadającą dwóm stanom danego układu, możemy wyznaczać dzięki temu, że energia wewnętrzna jest jednoznaczną funkcją stanu układu. Możemy to uzasadnić w następujący sposób.

Niech punkt *1* (rys. 18.3) przedstawia stan początkowy badanego układu. Od tego stanu początkowego możemy przejść do stanu *2*, np. przez ciąg stanów przejściowych, odpowiadających linii $1 \to a \to 2$. Od *2* możemy wrócić do *1* czy to przez ciąg stanów tych samych przebywanych w odwrotnej kolejności, a więc drogą $2 \to a \to 1$, czy też drogą inną (przez inne stany przejściowe), np. drogą $2 \to b \to 1$. Cała przemiana

$$1 \to a \to 2 \to b \to 1$$

jest przemianą kołową: stan końcowy nie różni się od początkowego, a zatem przy takiej przemianie wartość energii wewnętrznej układu nie ulega zmianie: $\Delta U = 0$. W różnych etapach tej przemiany mogły zachodzić zmiany cieplne (ciepło mogło być doprowadzane lub oddawane przez układ), mogła też być praca dostarczana układowi lub wykonywana przez układ, ale dla całego cyklu spełnione było równanie

$$\Delta Q = \Delta W.$$

Rozważmy przejście *1* → *a* → *2*. Tym razem stan początkowy i końcowy są różne. Odpowiada im różnica energii wewnętrznej ΔU. Wracając od stanu *2* do stanu *1* dowolną drogą przez *a*, *b* lub *c*, uzyskamy zmianę energii wewnętrznej taką samą co do wartości bezwzględnej, lecz przeciwną co do znaku, tzn. równą $-\Delta U$. Tylko wtedy bowiem przy przemianie kołowej *1* → *2* → *1* sumaryczne ΔU będzie równe zeru. Stąd wniosek, że zmiana energii wewnętrznej po przejściu od określonego stanu początkowego do określonego stanu końcowego nie zależy od rodzaju przejścia (od charakteru dokonanej przemiany), a zależy tylko od stanu początkowego i końcowego. Innymi słowy, *energia wewnętrzna jest jednoznaczną funkcją stanu układu.*

18.5. Pierwsza zasada termodynamiki

Zastanówmy się raz jeszcze nad przejściem od stanu *1* do stanu *2*. Przejściu temu odpowiada zmiana energii wewnętrznej ΔU układu. Ta zmiana energii wewnętrznej wiąże się z ciepłem ΔQ (doprowadzonym do układu lub oddanym przez układ) oraz z pracą ΔW (wydaną lub dostarczoną). Zapisujemy ten związek w postaci

$$\Delta Q = \Delta U + \Delta W, \tag{18.1}$$

co oznacza, że *kosztem ciepła ΔQ doprowadzonego do układu uzyskujemy wzrost jego energii wewnętrznej ΔU oraz pracę przez niego wykonaną ΔW.*

Równanie to wyraża w sposób ogólny treść *pierwszej zasady termodynamiki.* Ustalenie zmiany ΔU nie wymaga bliższego określania rodzaju przemiany, podczas której układ przeszedł od stanu *1* do stanu *2*, gdyż ΔU nie zależy od rodzaju przemiany. Nie można jednak tego powiedzieć z osobna o ΔQ i ΔW; mogą one ulegać zmianie w zależności od charakteru przemiany. Wielkości Q i W nie są funkcjami stanu układu, tzn. ciepło i praca nie charakteryzują stanu układu, lecz przemianę, jakiej podlega układ. Wymiana ciepła między układami, jak również praca wykonywana przez jeden układ na drugim, jest sposobem przekazywania energii. (Nawiązujemy tu wyraźnie do uwag wprowadzających do rozdz. 13.)

Każdą przemianę termodynamiczną można w myśli podzielić na nieskończenie wiele *przemian* tzw. *elementarnych*, tzn. przemian odnoszących się do dwóch stanów bardzo mało różniących się od siebie. Do takich przemian stosujemy I zasadę termodynamiki w postaci różniczkowej:

$$dQ = dU + dW^*. \tag{18.2}$$

* Trzeba tu podkreślić, że w (18.2) tylko dU stanowi różniczkę zupełną, gdyż — jak wyżej mówiliśmy — tylko zmiany energii wewnętrznej nie zależą od rodzaju przemiany. Można by to wyraźniej zaakcentować stosując następujący zapis zależności (18.2):

$$\delta Q = dU + \delta W.$$

Elementarną pracę związaną ze zmianą objętości ciała pod wpływem ciśnienia p możemy przedstawić, tak jak to robiliśmy w przypadku gazów:

$$dW = p dV \tag{18.3}$$

(praca ta jest dodatnia lub ujemna w zależności od tego, czy jest dostarczona układowi czy przez układ wydana).

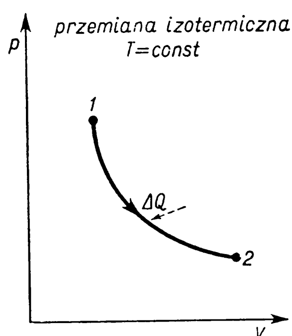
Od równań (18.1) lub (18.2), wyrażających w sposób ogólny I zasadę termodynamiki, łatwo można przejść do znanego nam już równania wyrażającego przemianę zamkniętą. W przemianie zamkniętej $\Delta U = 0$, a zatem

$$\Delta Q = \Delta W \quad \text{lub} \quad dQ = dW. \tag{18.4}$$

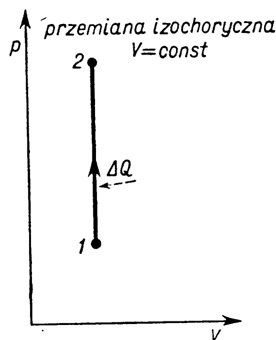
18.6. Zastosowanie I zasady termodynamiki do izoprzemian gazu doskonałego

Przez *izoprzemianę* rozumiemy taką przemianę, przy której jeden z parametrów charakteryzujących stan ciała pozostaje stały.

Zastosujmy I zasadę termodynamiki do różnych izoprzemian zachodzących w gazie doskonałym, pamiętając, że stosuje się ona do wszystkich ciał, niezależnie od ich stanu skupienia.



Rys. 18.4



Rys. 18.5

1. Przemiana izotermiczna (rys. 18.4)

Wiemy z doświadczenia Gay-Lussaca, że w przypadku gazu doskonałego energia wewnętrzna jest tylko funkcją temperatury — nie zależy od objętości i ciśnienia. Ponieważ w przemianie izotermicznej $T = \text{const}$, więc $U_1 = U_2$, czyli $dU = 0$. Stąd wniosek, że w elementarnej przemianie izotermicznej gazu doskonałego ciepło zamienia się całkowicie na pracę (lub odwrotnie) $dQ = dW$. W stanie *2* (rys. 18.4) objętość jest większa niż w stanie *1*, gaz rozszerzając się wykonał pracę ΔW. Izotermiczne rozszerzanie się wymagało doprowadzenia ciepła (zaznaczono to przerywaną strzałką na rysunku) w ilości $\Delta Q = \Delta W$ albo $\Delta Q = p\Delta V$ (dla przemiany elementarnej odpowiednio $dQ = p dV$).

2. Przemiana izochoryczna (rys. 18.5)

W elementarnej przemianie izochorycznej $dV = 0$, a zatem i $dW = 0$. Z I zasady termodynamiki otrzymujemy $dQ = dU$. Doprowadzenie ciepła do ciała przy niezmiennej objętości powoduje wzrost jego energii wewnętrznej równy dostarczonej energii cieplnej.

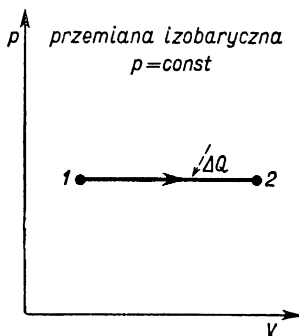
Ilość ciepła doprowadzona do 1 mola gazu doskonałego ogrzewanego w stałej objętości wyraża się iloczynem $C_V dT$, gdzie C_V jest ciepłem molowym w stałej objętości. Innymi słowy, *przyrost energii wewnętrznej 1 mola gazu doskonałego podczas ogrzania o dT wynosi*

$$dU = C_V dT. \tag{18.5}$$

W odniesieniu do dowolnej masy m gazu doskonałego zmiana energii wewnętrznej podczas ogrzania o dT wyraża się wzorem

$$dU = c_V m dT,$$

gdzie c_V jest ciepłem właściwym gazu w stałej objętości.



Rys. 18.6

3. Przemiana izobaryczna (rys. 18.6)

Zastosujmy znów nasze rozważania do jednego mola gazu doskonałego. Ogrzewamy jeden mol gazu doskonałego o dT pod stałym ciśnieniem. Ilość ciepła pobranego $dQ = C_p dT$, gdzie C_p — ciepło molowe pod stałym ciśnieniem. Przyrost energii wewnętrznej podczas ogrzania o dT wynosi $dU = C_V dT$. Gaz zwiększa swą objętość o dV_m, a więc praca wykonana $dW = p dV_m$.

Z I zasady termodynamiki wynika zatem, że

$$C_p dT = C_V dT + p dV_m.$$

Ale $pV_m = RT$, skąd jeśli $p = \text{const}$, to $p dV_m = R dT$. Po podstawieniu

$$C_p dT = C_V dT + R dT,$$

czyli

$$C_p - C_V = R.$$

W ten sposób powiązaliśmy stałą gazową z wartościami ciepła molowego C_p i C_V.

4. Przemiana adiabatyczna (rys. 18.7)

Jest to bardzo ważny typ przemiany, a mianowicie taka przemiana, podczas której nie ma wymiany ciepła z otoczeniem. Takie przemiany zachodzą np. w układach idealnie izolowanych cieplnie od otoczenia. Ponieważ w przemianie adiabatycznej $dQ = 0$, więc

$$dU = -dW.$$

Rozważmy adiabatyczne rozprężanie się jednego mola gazu doskonałego. W tej przemianie $dW = +p\,dV_m$. Praca zostaje wykonana kosztem energii wewnętrznej. Towarzyszy temu oziębienie gazu o dT, związane z ujemną zmianą energii wewnętrznej o $dU = C_V dT$. A zatem

$$C_V dT = -p\,dV_m. \tag{18.6}$$

Podczas tej przemiany ciśnienie ulega zmianie. Z równania Clapeyrona wynika, że

$$p = \frac{RT}{V_m}.$$

Po podstawieniu tej wartości do równania (18.6) otrzymujemy

$$C_V dT = -\frac{RT}{V_m} dV_m.$$

Po podzieleniu obu stron równania przez $C_V T$ otrzymujemy

$$\frac{dT}{T} = -\frac{R}{C_V}\frac{dV_m}{V_m}.$$

Po scałkowaniu tego równania otrzymujemy

$$\ln T = -\frac{R}{C_V}\ln V_m + \text{const}, \qquad \ln T + \ln V_m^{R/C_V} = \text{const}, \qquad \ln TV_m^{R/C_V} = \text{const},$$

skąd

$$TV_m^{R/C_V} = \text{const}. \tag{18.7}$$

Ale $R = C_p - C_V$, a zatem

$$\frac{R}{C_V} = \frac{C_p - C_V}{C_V} = \varkappa - 1,$$

gdzie

$$\varkappa = \frac{C_p}{C_V}.$$

Równanie (18.7) przyjmuje postać

$$TV_m^{\varkappa-1} = \text{const}. \tag{18.8}$$

Mnożąc je stronami przez $pV_m = RT$ otrzymujemy

$$TV_m^{\varkappa} p = \text{const}\, RT,$$

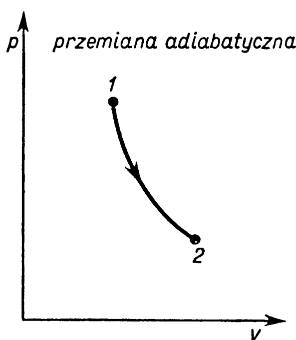
skąd

$$pV_m^{\varkappa} = \text{const}', \tag{18.9}$$

czyli

$$p_0 V_{0m}^{\varkappa} = p_1 V_{1m}^{\varkappa}. \tag{18.10}$$

Równania (18.8) i (18.9) wyrażają *prawo Poissona* dotyczące przemiany adiabatycznej 1 mola gazu doskonałego. Z ostatniego równania łatwo można się zorientować, że przebieg adiabaty w układzie współrzędnych (V, p) jest bardziej stromy niż przebieg izotermy (rys. 18.7), gdyż $\varkappa$ jest większe od jednostki. Odpowiada to oczywiście temu, że podczas adiabatycznych zmian objętości ciśnienie zmienia się gwałtowniej niż podczas przemian izotermicznych.



Rys. 18.7

Zestawiając wyniki dotyczące przemiany izotermicznej i adiabatycznej gazu doskonałego, stwierdzimy, że równanie izotermy $pV = \text{const}$ i równanie adiabaty $pV^{\varkappa} = \text{const}$ różnią się tylko wykładnikiem potęgi parametru V. W praktyce nie jest możliwe urzeczywistnienie warunków przemiany doskonale izotermicznej lub doskonale adiabatycznej. Pierwsza z nich wymagałaby doskonałej wymiany ciepła z otoczeniem, druga wymagałaby doskonałej izolacji cieplnej. W związku z niemożliwością stworzenia takich idealnych warunków przebieg procesów rzeczywistych odbiega od przemian izotermicznych i adiabatycznych. Jednym ze szczególnych przypadków możliwych przemian jest *proces politropowy*, którego równanie ma postać $pV^n = \text{const}$, z tym że wykładnik potęgi n, zwany wykładnikiem politropy, ma wartość zawartą między 1 (izoterma) i $\varkappa$ (adiabata).

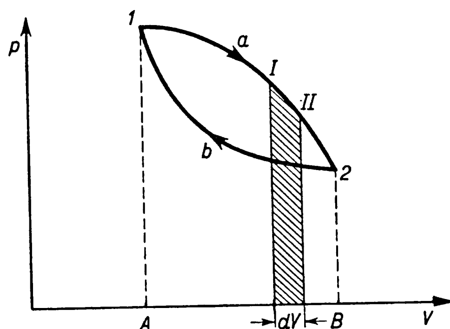
Znając prawo Poissona można łatwo obliczyć, jakie zmiany temperatury gazu towarzyszą określonym zmianom objętości. Okazuje się, że adiabatyczne sprężenie powietrza o temperaturze pokojowej do $\frac{1}{20}$ objętości pierwotnej wywołuje wzrost temperatury do około 700°C.

Znalazło to zastosowanie między innymi w silnikach Diesla, gdzie tłok gwałtownie spręża powietrze w cylindrze, powodując ogrzanie gazu do blisko 800°C. Wtryskiwana do cylindra mieszanka w tej temperaturze natychmiast się zapala, a powstające przy spalaniu gazy odrzucają tłok podtrzymując pracę silnika.

Powstawanie deszczu jest również wynikiem rozprężania adiabatycznego mas powietrza wznoszących się do góry i przechodzących z wyższego ciśnienia do niższego. Towarzyszy temu obniżenie temperatury, które może być tak znaczne, że para wodna znajdująca się w powietrzu z nienasyconej przekształca się w nasyconą i nadmiar jej ulega skropleniu.

18.7. Graficzne przedstawienie pracy

Wykresom przemian w układzie współrzędnych (V, p) można nadawać inną jeszcze interpretację. Krzywa $1 \to a \to 2$ (rys. 18.8) określa charakter przemiany, podczas której ciało przechodzi od stanu *1* do stanu *2*. Weźmy pod uwagę dwa punkty tej krzywej *I* i *II*, określające dwa stany bardzo mało różniące się od siebie. Przejście od stanu *I* do stanu *II* odpowiada elementarnej przemianie, podczas której zmiana objętości wynosi dV.



Rys. 18.8

Wiemy, że w czasie takiej przemiany układ wykonuje pracę równą pdV. Łatwo sprawdzić, że pole zakreskowanej na rys 18.8 figury o nieskończenie małej podstawie dV równa się też pdV. Przemianę skończoną od stanu *1* do stanu *2* można w myśli rozdzielić na nieskończenie wiele składowych elementarnych przemian. Pracę wykonaną podczas każdej z tych elementarnych przemian określać będzie pole odpowiedniej figury. Zsumowanie wszystkich pól wyznaczy powierzchnię zawartą między krzywą $1 \to a \to 2$ (określającą rodzaj przemiany), osią odciętych oraz prostymi *1A* i *2B*, równoległymi do osi rzędnych poprowadzonymi przez wartości V, odpowiadające stanom *1* i *2*. Pole wymienionego obszaru przedstawia pracę związaną z przemianą $1 \to a \to 2$. Przy liczbowej wartości tego pola napiszemy znak +, jeżeli, jak w przypadku zaznaczonym na rysunku, ciało rozprężając się wykonuje pracę.

Wróćmy od stanu *2* przez *b* do *1*. Analogiczne rozumowanie doprowadzi do wniosku, że pole figury $1 \to b \to 2 \to B \to A \to 1$ przedstawia pracę (ujemną) dostarczoną układowi podczas przemiany $2 \to b \to 1$. Pole $1 \to a \to 2 \to b \to 1$ ograniczone krzywą odpowiadającą przemianie kołowej (czyli cyklowi), wyraża nadwyżkę pracy wykonanej przez układ w pojedynczym cyklu nad pracą w tymże cyklu przez układ pobraną.

Jest to zatem efektywna praca ΔW wykonana przez układ w jednym cyklu. Ale wiemy, że w przemianie kołowej $\Delta Q = \Delta W$. Pole $1 \to a \to 2 \to b \to 1$ przedstawia więc równocześnie ΔQ, czyli nadwyżkę ciepła pobranego przez układ w jednym cyklu nad ciepłem w tymże cyklu oddanym na zewnątrz.

Dzięki takiej interpretacji wykresy przemian termodynamicznych zyskują nową wymowę.

18.8. Druga zasada termodynamiki

Zanim przejdziemy do wysłowienia drugiej zasady termodynamiki, zajmiemy się urządzeniami przetwarzającymi ciepło na pracę. Urządzenia takie nazywamy silnikami termodynamicznymi. Silnik *termodynamiczny* (prosty) jest to urządzenie zasilane energią cieplną, pracujące okresowo w zamkniętych obiegach i dostarczające pracę w każdym obiegu bez zużywania się ciała roboczego, czyli czynnika termodynamicznego.

Druga zasada termodynamiki określa możliwości przemiany ciepła na pracę, czyli, innymi słowy, podaje warunki pracy silnika termodynamicznego.

Treść drugiej zasady termodynamiki ujmujemy w następujący sposób: *zamiana ciepła na pracę w silniku termodynamicznym jest możliwa jedynie wtedy, gdy źródło dostarczające ciepła ma temperaturę wyższą od najzimniejszego ciała w jego otoczeniu.*

Pierwszą zasadę termodynamiki można traktować jako rozszerzenie zasady zachowania energii na zjawiska cieplne. Omawiając zasadę zachowania energii podkreśla się zazwyczaj fakt niemożliwości zbudowania perpetuum mobile pierwszego rodzaju. *Perpetuum mobile pierwszego rodzaju* byłaby to maszyna, która bez zasilania jej pracą lub energią stale wykonywałaby pracę. Wnioskiem wypływającym z drugiej zasady termodynamiki jest stwierdzenie niemożliwości zbudowania perpetuum mobile drugiego rodzaju.

Perpetuum mobile drugiego rodzaju byłby to silnik termodynamiczny, który stale przetwarzałby ciepło na pracę nawet w przypadku braku różnicy temperatur między źródłem ciepła i otoczeniem. Praca silnika termodynamicznego (prostego) może przebiegać tylko w takich warunkach, gdy ciepło jest pobierane ze źródła o temperaturze wyższej, zwanego w technice *górnym źródłem ciepła*. Część tego ciepła zostaje przetworzona na pracę, a reszta jest oddawana otoczeniu o temperaturze niższej (zwanemu *dolnym źródłem ciepła*).

18.9. Sprawność silników termodynamicznych. Cykl Carnota

Jednym z pierwszoplanowych zagadnień z dziedziny termodynamiki jest zagadnienie sprawności silników termodynamicznych. Przez *sprawność η silnika termodynamicznego* rozumiemy stosunek efektywnej pracy W wykonanej przez silnik w pojedynczym cyklu do ciepła Q_1 dostarczonego silnikowi w tym cyklu:

$$\eta = \frac{W}{Q_1}. \tag{18.11}$$

Zarówno rozważania teoretyczne, jak i pomiary wykazują, że sprawność silnika

termodynamicznego pracującego między temperaturami T_1 (np. temperatura kotła w maszynie parowej) i T_2 (np. temperatura chłodnicy) zależy w dużym stopniu od tego, czy silnik termodynamiczny pracuje w sposób odwracalny czy nieodwracalny.

Przez *przemianę odwracalną* rozumiemy taką przemianę, w której kierunek przebiegu w każdej chwili może się dowolnie zmieniać.

Przemiana odwracalna składa się z ciągu stanów, z których każdy różni się nieskończenie mało od stanu równowagi. Prędkość takiej przemiany musi być nieskończenie mała, gdyż w przeciwnym przypadku — przy istnieniu bodźców zewnętrznych, nadających zjawisku skończoną prędkość przebiegu — kierunek przemiany nie mógłby się dowolnie zmieniać.

Należy sobie zdawać sprawę, że wprowadzenie pojęcia przemian odwracalnych jest znów pewną idealizacją zjawisk. Zjawiska zachodzące samorzutnie w naszym otoczeniu są nieodwracalne. Możemy jednak sztucznie stwarzać warunki takie, że zachodzące przemiany niesłychanie mało różnią się od przemian odwracalnych.

Dla uzyskania lepszego obrazu przemian odwracalnych i nieodwracalnych rozważmy kilka przykładów.

1. Gaz zamknięty w naczyniu pod obciążonym tłokiem jest silnie sprężony. Po usunięciu z powierzchni tłoka odważników gaz gwałtownie się rozpręża ze względu na przewagę prężności gazu nad ciśnieniem atmosferycznym. Zjawisko przebiega ze skończona prędkością; kierunek zjawiska w dowolnej chwili nie może zmienić się na przeciwny. Sprężony gaz nie zacznie samorzutnie sprężać się w jeszcze większym stopniu. Omawiana przemiana jest nieodwracalna.

2. Rozprężanie się gazu pod ciśnieniem zewnętrznym równym ciśnieniu wewnętrznemu jest przykładem przemiany odwracalnej. Przy zastosowaniu nieskończenie małej nadwyżki ciśnienia zewnętrznego nad ciśnieniem wewnętrznym gaz się spręża, a przy zastosowaniu nieskończenie małej nadwyżki ciśnienia wewnętrznego nad zewnętrznym — gaz się rozpręża.

3. Mieszanina wody z lodem umieszczona jest w środowisku o temperaturze 0°C. Nieskończenie małe obniżenie temperatury powoduje krzepnięcie wody. Nieskończenie małe podniesienie temperatury powoduje topnienie lodu. Przemiana jest odwracalna.

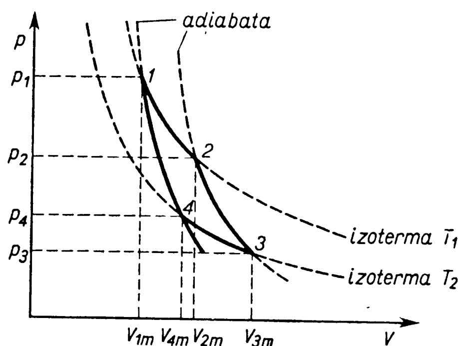
Można udowodnić, że *sprawność η_0 silnika termodynamicznego, pracującego w przemianach kołowych odwracalnych, jest zawsze większa od sprawności η silnika pracującego w przemianach kołowych nieodwracalnych przy tej samej różnicy temperatur kotła i chłodnicy*:

$$\eta_0 > \eta. \tag{18.12}$$

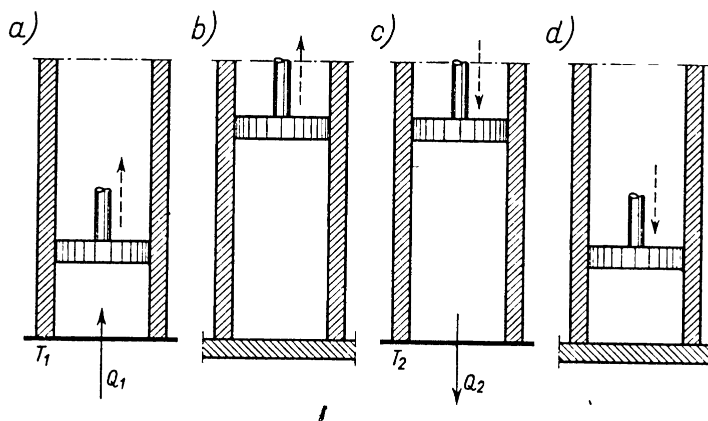
Można też udowodnić, że *maksymalna sprawność różnych silników odwracalnych pracujących między tymi samymi temperaturami nie zależy od rodzaju ciała roboczego.*

Wyprowadzimy wzór na sprawność silnika termodynamicznego odwracalnego, w którym ciałem roboczym, czyli czynnikiem termodynamicznym, jest gaz doskonały. Rozważana przemiana kołowa składać się będzie z dwóch przemian odwracalnych izotermicznych, odpowiadających temperaturom T_1 i T_2, oraz z dwóch przemian odwracalnych adiabatycznych. Jest to tzw. *cykl Carnota*.

Rysunek 18.9 przedstawia cykl Carnota w układzie współrzędnych (V, p). Przemiana $1 \rightarrow 2$ odpowiada izotermicznemu rozprężaniu się gazu doskonałego w temperaturze T_1 (temperatura kotła). Przejście $2 \rightarrow 3$ odbywa się wzdłuż adiabaty. Gaz rozprężając się adiabatycznie oziębia się do temperatury T_2. Przemiana $3 \rightarrow 4$ związana jest z izotermicznym sprężaniem się gazu w temperaturze T_2 (temperatura chłodnicy). Powrót do stanu wyjściowego odpowiada adiabatycznemu sprężeniu gazu. Jeszcze raz podkreślamy, że wszystkie cztery przemiany mają przebieg odwracalny.



Rys. 18.9



Rys. 18.10

Opisane przemiany można przedstawić schematycznie w następujący sposób. Wyobraźmy sobie cylinder roboczy zamknięty tłokiem (rys. 18.10). Ściany boczne cylindra i tłok są nieprzepuszczalne dla ciepła (doskonale izolują pod względem cieplnym, czyli stanowią tzw. *osłony adiabatyczne*). Cylinder roboczy można dowolnie łączyć z kotłem o dużej pojemności cieplnej lub z chłodnicą również o dużej pojemności cieplnej. Duża pojemność cieplna kotła zapewnia stałość temperatury T_1 nawet wtedy, gdy kocioł dostarcza ciału roboczemu pewnej ilości ciepła. Podobnie duża pojemność cieplna chłodnicy zapewnia stałość temperatury T_2 nawet wtedy, gdy ciało robocze oddaje chłodnicy pewną ilość ciepła.

W przejściowych stadiach cyklu, gdy cylinder roboczy nie jest połączony ani z kotłem, ani z chłodnicą, dno jego zostaje również pokryte osłoną adiabatyczną. W tym czasie w silniku termodynamicznym dokonują się przemiany adiabatyczne. Rys. 18.10 przedstawia kolejne części cyklu. Podwójne ścianki oznaczają osłony adiabatyczne.

W pierwszej części cyklu (rys. 18.10a) podczas izotermicznego rozprężania się gazu zostaje wykonana praca W_1 kosztem ciepła Q_1 dostarczonego z kotła. Strzałka ciągła wskazuje kierunek przepływu ciepła, strzałka przerywana — kierunek ruchu tłoka. Zgodnie z I zasadą termodynamiki $W_1 = Q_1$.

W drugiej części cyklu (rys. 18.10b) cylinder roboczy jest całkowicie odcięty od otoczenia osłonami adiabatycznymi. Gaz w dalszym ciągu rozszerza się, ale tym razem adiabatycznie. W tej części przemiany ciepło nie jest ani dostarczane, ani oddawane, natomiast praca W_2 jest wykonana przez układ. Zgodnie z I zasadą termodynamiki obowiązuje tu zależność $W_2 = -\Delta U$.

W trzeciej części cyklu (rys. 18.10c) cylinder roboczy jest połączony z chłodnicą. Gaz jest sprężany izotermicznie w temperaturze T_2. Tym razem ciepło Q_2 jest oddawane chłodnicy, a praca W_3 wykonywana na układzie.

W czwartej części cyklu (rys. 18.10d) ścianki cylindra stanowią znowu osłony adiabatyczne. Gaz jest adiabatycznie sprężany aż do osiągnięcia temperatury T_1 i objętości odpowiadającej stanowi początkowemu. Praca W_4 jest przy tym dostarczana układowi. Żadnej wymiany ciepła z otoczeniem w tej części cyklu nie ma.

W przemianie kołowej obowiązuje $\Delta Q = \Delta W$, czyli

$$Q_1 - Q_2 = W,$$

gdzie

$$W = W_1 + W_2 + W_3 + W_4,$$

przy czym W_1 i W_2 są dodatnie, a W_3 i W_4 są ujemne.

Dla wyznaczenia sprawności silnika ze wzoru $\eta_0 = W/Q_1$ musimy obliczyć W_1, W_2, W_3 i W_4.

Dla uproszczenia zakładamy, że ciałem roboczym jest jeden mol gazu doskonałego. Przemianie izotermicznej zachodzącej w temperaturze T odpowiada praca

$$W_1 = \int_{V_{1m}}^{V_{2m}} p\,dV_m.$$

Podstawiając

$$p = \frac{RT_1}{V_m},$$

otrzymamy

$$W_1 = RT_1 \int_{V_{1m}}^{V_{2m}} \frac{dV_m}{V_m} = RT_1(\ln V_{2m} - \ln V_{1m}),$$

czyli

$$W_1 = RT_1 \ln\frac{V_{2m}}{V_{1m}} \text{ (praca dodatnia).} \qquad (18.13)$$

W przemianie adiabatycznej temperatura zmienia się od T_1 do T_2, a zatem praca

$$W_2 = -\int_{T_1}^{T_2} C_V dT = C_V(T_1 - T_2) \text{ (praca dodatnia).} \tag{18.14}$$

Dla przemiany izotermicznej zachodzącej w temperaturze T_2, przez analogię do wzoru na W_1 piszemy

$$W_3 = RT_2 \ln \frac{V_{4m}}{V_{3m}} \text{ (praca ujemna).}$$

Dla przemiany adiabatycznej, której odpowiada zmiana temperatur od T_2 do T_1, otrzymujemy

$$W_4 = C_V(T_2 - T_1) \text{ (praca ujemna).}$$

Przy sumowaniu prace W_2 i W_4 redukują się, a zatem

$$W = RT_1 \ln \frac{V_{2m}}{V_{1m}} + RT_2 \ln \frac{V_{4m}}{V_{3m}}.$$

Korzystając z równania Poissona udowodnimy, że

$$-\ln \frac{V_{2m}}{V_{1m}} = \ln \frac{V_{4m}}{V_{3m}}.$$

Wypisujemy równania Poissona dla przemian adiabatycznych od objętości V_{4m} w temperaturze T_2 do objętości V_{1m} w temperaturze T_1 oraz od objętości V_{2m} w temperaturze T_1 do objętości V_{3m} w temperaturze T_2

$$T_1 V_{1m}^{\varkappa-1} = T_2 V_{4m}^{\varkappa-1},$$

$$T_1 V_{2m}^{\varkappa-1} = T_2 V_{3m}^{\varkappa-1}.$$

Dzielimy stronami powyższe równania i wyciągamy pierwiastek stopnia $\varkappa - 1$:

$$\frac{V_{1m}}{V_{2m}} = \frac{V_{4m}}{V_{3m}}.$$

Logarytmując to wyrażenie otrzymujemy

$$\ln \frac{V_{1m}}{V_{2m}} = \ln \frac{V_{4m}}{V_{3m}}, \qquad -\ln \frac{V_{2m}}{V_{1m}} = \ln \frac{V_{4m}}{V_{3m}},$$

a zatem

$$W = RT_1 \ln \frac{V_{2m}}{V_{1m}} - RT_2 \ln \frac{V_{2m}}{V_{1m}},$$

$$W = R(T_1 - T_2) \ln \frac{V_{2m}}{V_{1m}}.$$

Pamiętając, że praca wykonana podczas pierwszej przemiany izotermicznej równa się ilości pobranego ciepła Q_1, znajdziemy że

$$\eta_0 = \frac{W}{Q_1} = \frac{R(T_1 - T_2) \ln \frac{V_{2m}}{V_{1m}}}{RT_1 \ln \frac{V_{2m}}{V_{1m}}},$$

skąd

$$\eta_0 = \frac{T_1 - T_2}{T_1}. \tag{18.15}$$

Można udowodnić, że wyprowadzony wzór (18.15) na sprawność procesu Carnota podaje teoretyczną maksymalną sprawność silnika termodynamicznego, pracującego między stałymi temperaturami T_1 i T_2,

$$\eta_{0\,\mathrm{Carnota}} = \eta_{0\,\max}.$$

Gdyby temperatura chłodnicy T_2 była równa zeru, to $\eta_{\max} = 1$, czyli wynosiłoby 100%, gdyby zaś nie było różnicy temperatur między kotłem i chłodnicą ($T_1 = T_2$), to $\eta_{\max} = 0$. Zgodnie z drugą zasadą termodynamiki silnik taki nie mógłby pracować. Ustalenie wzoru na sprawność maksymalną przyczyniło się do skierowania wysiłków mających na celu ulepszenie silników termodynamicznych na właściwe tory. Nie wybór ciała roboczego, lecz zwiększenie różnicy temperatur kotła i chłodnicy ma zasadnicze znaczenie. Nigdy jednak η_{rzecz} nie osiąga wartości teoretycznej $\eta_{\max}$. Dla ilustracji podajemy kilka liczb: współczesna maszyna parowa pracująca w temperaturach między 250 i 20°C ma sprawność 18% zamiast teoretycznych 44%. Turbiny parowe osiągają sprawność do czterdziestu kilku procentów. Silniki spalinowe, benzynowe i Diesla mają sprawność do 40%. Większą jeszcze sprawność mają silniki odrzutowe.

Jest rzeczą jasną, że w technice dąży się do budowy silników termodynamicznych o sprawności możliwie zbliżonej do sprawności maksymalnej. Osiągnięcie sprawności maksymalnej jest jednak niemożliwe, gdyż m. in. wymagałoby stosowania procesów odwracalnych, o przebiegu niesłychanie powolnym, co jest oczywiście nie do przyjęcia ze względów praktycznych.

Łatwo stwierdzić, że sprawność dowolnego, niekoniecznie odwracalnego silnika termodynamicznego wyraża się wzorem

$$\eta = \frac{Q_1 - Q_2}{Q_1}. \tag{18.16}$$

Wynika to stąd, że w każdej przemianie kołowej $W = Q_1 - Q_2$.

Warto też zwrócić uwagę, że obiegi stosowane w technice są nie tylko nieodwracalne, ale także niezamknięte, gdyż czynnik roboczy (np. para wodna, spalona mieszanka itp.) jest usuwany na zewnątrz w końcowej fazie cyklu.

P y t a n i e. Jak zmienią się wzory (18.13) i (18.14) wyrażające pracę izotermicznego i adiabatycznego rozprężania się 1 mola gazu doskonałego, gdy odniesiemy je do masy m kg gazu doskonałego?

18.10. Termodynamiczna skala temperatur

Zestawmy pewne dane dotyczące pracy silników termodynamicznych.

1. Sprawność η każdego silnika termodynamicznego wynosi

$$\eta = \frac{Q_1 - Q_2}{Q_1}.$$

2. Maksymalna sprawność $\eta_{0\,max}$ silnika termodynamicznego odwracalnego, pracującego w obiegu Carnota między stałymi temperaturami kotła i chłodnicy, zależy tylko od tych temperatur:

$$\eta_{0\,max} = \frac{T_1 - T_2}{T_1}.$$

3. Sprawność silnika termodynamicznego odwracalnego nie zależy od rodzaju ciała roboczego.

Wyobraźmy sobie silnik termodynamiczny, pracujący odwracalnie między temperaturami T_1 i T_2 i pobierający w każdym cyklu ilość ciepła Q_1 ze źródła o temperaturze T_1. Ilość ciepła Q_2 oddawana chłodnicy w każdym cyklu jest w tych warunkach ściśle określona. Gdybyśmy zachowując stałą temperaturę T_1 i ilość ciepła pobieranego Q_1 zmieniali temperaturę T_2, to i wartość Q_2 ulegałaby zmianie. Stąd wniosek, że wielkości Q_1, Q_2 oraz T_1 i T_2 są ze sobą związane. Można by wprowadzić umowę, że temperatury kotła i chłodnicy będą się wyrażały tymi samymi liczbami, jakimi wyrażają się odpowiadające im ilości ciepła Q_1 i Q_2.

Kelvin zaproponował inną umowę, a mianowicie takie dobranie liczb określających temperaturę kotła i chłodnicy w nowej skali (oznaczmy je przez T_1^x i T_2^x), aby spełniona była zależność

$$\frac{Q_1}{Q_2} = \frac{T_1^x}{T_2^x}.$$

Stąd wynika, że

$$1 - \frac{Q_2}{Q_1} = 1 - \frac{T_2^x}{T_1^x},$$

co można inaczej przedstawić jako

$$\frac{Q_1 - Q_2}{Q_1} = \frac{T_1^x - T_2^x}{T_1^x}.$$

Warunek

$$\frac{Q_1}{Q_2} = \frac{T_1^x}{T_2^x}$$

nie określa jednoznacznie temperatury w nowej skali. Dodatkowo ustalamy w nowej skali różnicę temperatur wrzenia wody pod ciśnieniem normalnym i topnienia lodu pod ciśnieniem normalnym, zakładając, że

$$T^x_{\text{wrzenia wody}} - T^x_{\text{topnienia lodu}} = 100.$$

Uwzględniając oba te warunki można obliczyć, jaką liczbą wyraża się np. temperatura wrzenia wody w nowej skali. Wykorzystujemy do tego celu znaną liczbowo sprawność silnika termodynamicznego odwracalnego, pracującego między temperaturą wrzenia wody i topnienia lodu:

$$\eta_{0(T_{wrz} \to T_{topn})} = 0{,}268.$$

Stąd

$$0{,}268 = \frac{100}{T^x_{\text{wrz wody}}}, \quad \text{czyli} \quad T^x_{\text{wrz wody}} = 373{,}15.$$

Skalę w ten sposób określoną, zwaną skalą termodynamiczną lub skalą Kelvina, nazywa się też często skalą bezwzględną. Ma to swoje uzasadnienie w tym, że definicja temperatury w tej skali jest niezależna od rodzaju ciała termodynamicznego, czyli ciała roboczego w silniku termodynamicznym. Przypominamy, że

$$T\text{K} = (t+273{,}15)^\circ\text{C}.$$

Dotychczasowe rozważania związane z działaniem silników termodynamicznych dotyczyły silników termodynamicznych prostych, tzn. przetwarzających ciepło na pracę. Drugim rodzajem silników termodynamicznych są silniki termodynamiczne odwrotne, które mogą działać jako chłodziarki lub jako pompy cieplne.

Jeśli w czasie każdego cyklu pracy silnik pobiera ciepło Q_2 z chłodnicy o temperaturze T_2, a oddaje ciepło $Q_1(Q_1 > Q_2)$ źródłu o temperaturze $T_1(T_1 > T_2)$, to brakującą różnicę Q_1-Q_2 może pokryć albo dostarczona z zewnątrz praca, albo ciepło doprowadzone z dodatkowego źródła. W pierwszym przypadku silnik termodynamiczny może działać jako *chłodziarka*, tzw. *sprężarkowa*, obniżająca temperaturę chłodnicy, albo jako *pompa cieplna* dostarczająca źródłu dodatkowych ilości ciepła kosztem pobranej z zewnątrz pracy. W drugim przypadku, gdy różnicę Q_1-Q_2 pokrywa nie praca, lecz ciepło doprowadzane z dodatkowego źródła, mówimy o *chłodziarkach absorpcyjnych*.

W chłodziarkach sprężarkowych jako czynnik roboczy stosowane są najczęściej pary amoniaku, dwutlenku węgla, a ostatnio coraz szerzej wchodzą w użycie tzw. freony będące fluorochlorowymi pochodnymi niższych węglowodorów. W chłodziarkach absorpcyjnych czynnikiem roboczym jest wodny roztwór amoniaku.

18.11. Zasada degradacji energii

Zastanówmy się nad tym, co decyduje o walorach użytkowych energii: czy tylko jej ilość, czy też i jakość, tzn. rodzaj energii i warunki, w jakich występuje. Wody mórz i oceanów przedstawiają olbrzymie zasoby energii. Oziębienie tych mas wód o ułamki stopnia związane byłoby z wydzieleniem bardzo wielkich ilości ciepła. Byłaby to jednak energia mało wartościowa pod względem ekonomicznym, gdyż przekształcenie jej np. na pracę wymagałoby stworzenia takich warunków, aby otoczenie tych wód miało temperaturę od nich niższą. Tylko wtedy to „otoczenie" mogłoby stanowić chłodnicę silnika termodynamicznego.

Wszystkie zjawiska zachodzące samorzutnie w przyrodzie są zjawiskami nieodwracalnymi. Mają one określony kierunek przebiegu, a mianowicie zawsze taki, że w czasie trwania takich zjawisk energia określonego rodzaju, np. energia mechaniczna, chemiczna, elektryczna, magnetyczna, przetwarza się na ciepło. Tym przemianom może towarzyszyć powstanie pewnych różnic temperatur. Wiemy jednak, że jeżeli obok siebie istnieją dwa ciała o różnych temperaturach, to samorzutny przepływ ciepła od ciała o temperaturze wyższej do ciała o temperaturze niższej powoduje wyrównanie się temperatur. Po wyrównaniu się temperatur ciał niemożliwa już jest przemiana ciepła na pracę. Nawet w idealnym procesie odwracalnym silnika termodynamicznego z ogól-

nej ilości ciepła Q_1, dostarczonej przez kocioł, na pracę przekształca się tylko część ciepła, reszta zostaje oddana chłodnicy. Ta reszta przedstawia zasób energii trudniejszy już do przetworzenia na pracę lub inny rodzaj energii. W czasie przemiany np. na pracę ta ilość ciepła wymagałaby użycia nowej chłodnicy o jeszcze niższej temperaturze.

Ta kierunkowość przemian w przyrodzie, objawiająca się w przetwarzaniu się pracy lub jakiejkolwiek energii na ciepło, odpowiada zasadzie degradacji albo rozpraszania się energii.

Nawiązując do pierwszej zasady termodynamiki powiemy: *w układzie odosobnionym ogólna ilość zasobów energii jest stała, lecz zjawiska zachodzące samorzutnie w takim układzie prowadzą do zmniejszania się jej wartości użytkowej.*

18.12. Nierówność Clausiusa

Omówioną wyżej kierunkowość przebiegu zjawisk w przyrodzie można ująć ilościowo za pomocą nowej funkcji stanu układu, zwanej entropią. Zanim przejdziemy do określenia entropii musimy się zająć tzw. nierównością Clausiusa.

Wiemy, że sprawność silnika termodynamicznego

$$\eta = \frac{Q_1 - Q_2}{Q_1}.$$

Dla silnika odwracalnego, pracującego w obiegu Carnota między stałymi temperaturami kotła i chłodnicy

$$\eta_{0\,\max} = \frac{T_1 - T_2}{T_1}.$$

Sprawność silnika pracującego między temperaturami T_1 i T_2 może być co najwyżej równa $\eta_{0\,\max}$, czyli

$$\frac{Q_1 - Q_2}{Q_1} \leqslant \frac{T_1 - T_2}{T_1}.$$

Stąd

$$1 - \frac{Q_2}{Q_1} \leqslant 1 - \frac{T_2}{T_1},$$

a zatem

$$\frac{Q_2}{Q_1} \geqslant \frac{T_2}{T_1}.$$

Mnożąc przez wyrażenie dodatnie Q_1/T_2 otrzymujemy

$$\frac{Q_2}{T_2} \geqslant \frac{Q_1}{T_1},$$

czyli

$$0 \geqslant \frac{Q_1}{T_1} - \frac{Q_2}{T_2}.$$

Można tę nierówność przepisać w postaci

$$\frac{Q_1}{T_1}+\frac{Q_2}{T_2}\leqslant 0$$

pamiętając, że Q_2 jest wielkością ujemną. Znak równości obowiązuje dla odwracalnego obiegu Carnota.

Można udowodnić, że analogiczna zależność może być rozszerzona na większą liczbę przemian tworzących łącznie cykl zamknięty. Niech ilości ciepła odpowiadające poszczególnym częściom cyklu wynoszą ΔQ_1, ΔQ_2, ΔQ_3, ..., ΔQ_n i będą pobierane lub oddawane (a więc dodatnie lub ujemne) odpowiednio w temperaturach T_1, T_2, ..., T_n. Wtedy

$$\frac{\Delta Q_1}{T_1}+\frac{\Delta Q_2}{T_2}+\frac{\Delta Q_3}{T_3}+\ldots+\frac{\Delta Q_n}{T_n}\leqslant 0. \tag{18.17}$$

Jest to tzw. *nierówność Clausiusa. Znak równości obowiązuje przy przemianach odwracalnych, znak nierówności — przy przemianach nieodwracalnych.* Stosunek $\Delta Q/T$ — ciepła pobranego do temperatury, w jakiej jest ono pobierane (lub oddawane) — nazywamy *ciepłem zredukowanym.* Przez temperaturę, w której ciepło jest pobierane, rozumiemy temperaturę źródła dostarczającego ciepło.

Z nierówności Clausiusa wynika, że *suma wartości ciepła zredukowanego w przemianie odwracalnej kołowej równa się zeru, a w przemianie kołowej nieodwracalnej jest mniejsza od zera.*

18.13. Entropia

Opierając się na równości

$$\frac{\Delta Q_1}{T_1}+\frac{\Delta Q_2}{T_2}+\frac{\Delta Q_3}{T_3}+\ldots+\frac{\Delta Q_n}{T_n}=0,$$

obowiązującej w przemianach kołowych odwracalnych, przeprowadzamy rozumowanie następujące.

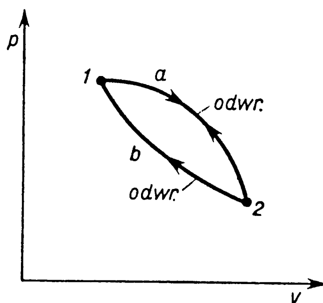
Podczas przejścia odwracalnego *1 → a → 2* (rys. 18.11) suma ilości ciepła zredukowanego $\sum_1^2 \Delta Q/T$ zmieniła się o pewną wartość. Oznaczmy tę zmianę przez ΔS. Zmiana ciepła zredukowanego podczas przejścia odwracalnego *2 → a → 1*, czyli $\sum_2^1 \Delta Q/T$ ma przeciwny znak, lecz wartość tę samą, gdyż w przemianie odwracalnej kołowej łączna wartość ciepła zredukowanego równa się zeru. Od stanu *2* do stanu *1* moglibyśmy wrócić za pośrednictwem innej odwracalnej przemiany i znowu łączna wartość ciepła zredukowanego musiałaby być równa zeru. Innymi słowy, *łączne ciepło zredukowane, odpowiadające przejściu odwracalnemu od stanu 1 do stanu 2, nie zależy od rodzaju przemiany, a zależy od stanu początkowego i końcowego.* Wprowadźmy zatem nową funkcję

stanu, zwaną *entropią*, której zmiana ΔS, odpowiadająca przejściu odwracalnemu od stanu *1* do stanu *2*, wyraża się wzorem

$$\Delta S = S_2 - S_1,$$

a równocześnie

$$S_2 - S_1 = \sum_1^2 \frac{\Delta Q}{T}. \qquad (18.18)$$



Rys. 18.11

Entropia danego stanu, podobnie jak energia wewnętrzna, jest określana tylko w odniesieniu do entropii wybranego stanu początkowego „zerowego", któremu umownie przypisujemy entropię S_0. Na przykład, badanemu stanowi (pierwszemu) odpowiada entropia S_1, przy czym

$$S_1 - S_0 = \sum_0^1 \frac{\Delta Q}{T},$$

czyli

$$S_1 = S_0 + \sum_0^1 \frac{\Delta Q}{T}. \qquad (18.19)$$

Entropia stanu badanego wyraża się więc sumą entropii stanu zerowego i łącznego ciepła zredukowanego, odpowiadającego przejściu odwracalnemu od stanu zerowego do stanu badanego. Podkreślamy, że w definicji entropii jest wyraźnie mowa o przemianie odwracalnej.

Dla elementarnej przemiany odwracalnej możemy napisać

$$\frac{dQ}{T} = dS,$$

czyli

$$dQ = TdS,$$

co oznacza, że *ciepło dostarczone czynnikowi w elementarnej przemianie odwracalnej wyraża się iloczynem temperatury bezwzględnej i elementarnego przyrostu entropii charakteryzującego tę elementarną przemianę.*

Dotychczas braliśmy pod uwagę tylko entropię czynnika termodynamicznego. Mo-

żemy pojęcie entropii stosować i do źródła ciepła i do układu ciał biorących udział w przemianach.

Załóżmy, że rozważamy układ złożony z czynnika termodynamicznego i źródła, będący układem odosobnionym (wystarcza nawet założenie, że układ jest adiabatycznie izolowany od otoczenia). Wtedy wymiana ciepła zachodzi tylko między źródłem i czynnikiem termodynamicznym: ile ciepła oddaje źródło (ΔQ ujemne), tyle pobiera czynnik (ΔQ dodatnie). Gdy zakładamy odwracalność przemiany, to temperatury źródła i czynnika muszą być jednakowe. A zatem zmiana entropii S_c czynnika termodynamicznego przy odwracalnym przejściu od stanu *1* do *2* wynosi

$$S_{c2}-S_{c1} = \sum_{1}^{2} \frac{\Delta Q}{T}.$$

Analogicznie, dla źródła zmiana S_z wynosi

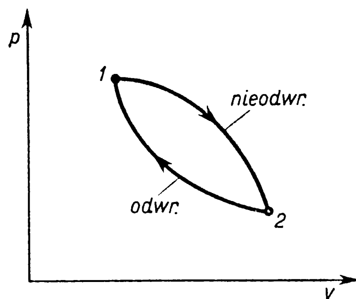
$$S_{z2}-S_{z1} = \sum_{1}^{2} \frac{\Delta Q}{T}.$$

Prawe strony są równe co do wartości, lecz przeciwne co do znaku, a więc

$$S_{c2}-S_{c1}+S_{z2}-S_{z1} = 0,$$

$$S_{c1}+S_{z1} = S_{c2}+S_{z2}.$$

Innymi słowy, *w przemianie odwracalnej entropia całkowita układu odosobnionego nie ulega zmianie.*



Rys. 18.12

Aby określić za pomocą entropii kierunkowość przebiegu zjawisk zachodzących samorzutnie w przyrodzie, rozważmy w układzie odosobnionym taką przemianę, w której przejście od stanu *1* do *2* odbywa się w sposób nieodwracalny, a przejście od stanu *2* do *1* — w sposób odwracalny (rys. 18.12). Całość obiegu musimy zatem traktować jako zjawisko nieodwracalne, czyli dla badanego obiegu obowiązuje nierówność Clausiusa

$$\sum_{1\,\text{nieodwr}}^{2} \frac{\Delta Q}{T} + \sum_{2\,\text{odwr}}^{1} \frac{\Delta Q}{T} < 0.$$

Sumę wartości ciepła zredukowanego w przemianie odwracalnej od stanu *2* do *1* można zastąpić różnicą entropii układu $S_1 - S_2$ (tym razem S_1 oznacza entropię stanu końcowego).

$$\sum_{1\,\text{nieodwr}}^{2} \frac{\Delta Q}{T} + S_1 - S_2 < 0,$$

skąd

$$S_2 - S_1 > \sum_{1\,\text{nieodwr}}^{2} \frac{\Delta Q}{T}.$$

Ale w układzie odosobnionym (izolowanym cieplnie) $\sum_{1\,\text{nieodwr}}^{2} \frac{\Delta Q}{T} = 0$, gdyż układ jako całość nie pobiera i nie oddaje ciepła, czyli

$$S_2 - S_1 > 0. \tag{18.20}$$

Stąd

$$S_2 > S_1, \tag{18.21}$$

a zatem w *układzie odosobnionym* (*adiabatycznie izolowanym*) *zachodzą samorzutnie takie zjawiska nieodwracalne, podczas których entropia rośnie.* W tej zasadzie wzrostu entropii jest właśnie zawarte *kryterium kierunkowości przemian w przyrodzie.*

Celem bliższego zapoznania się z pojęciem entropii rozważymy cztery przypadki.

1. Jaka jest zmiana entropii podczas odwracalnego, izotermicznego rozprężania się jednego mola gazu doskonałego? Wiemy, że praca wykonana podczas takiej przemiany w temperaturze T przy zmianie objętości od V_{1m} do V_{2m} wyraża się wzorem

$$W = RT \ln \frac{V_{2m}}{V_{1m}}.$$

To samo wyrażenie przedstawia ilość ciepła Q pobraną w tej przemianie w temperaturze T:

$$Q = RT \ln \frac{V_{2m}}{Q_{1m}}.$$

Pamiętając, że w przemianie odwracalnej ciepło zredukowane równe jest różnicy entropii stanu końcowego i stanu początkowego, otrzymujemy dla przemiany izotermicznej

$$S_2 - S_1 = R \ln \frac{V_{2m}}{V_{1m}}.$$

2. Sprawdzić, czy samorzutnemu przekazywaniu ciepła przez ciało A o temperaturze wyższej (T_1) ciału B o temperaturze niższej (T_2) towarzyszy wzrost entropii? Niech oba ciała stanowią układ izolowany cieplnie od otoczenia. Innymi słowy, ilość ciepła oddana przez ciało A równa się ilości ciepła pobranego przez ciało B. W tej nieodwracalnej przemianie

$$S_2 - S_1 > \sum_{1}^{2} \frac{\Delta Q}{T},$$

czyli
$$S_2-S_1 > -\frac{Q}{T_1}+\frac{Q}{T_2}.$$

Przy Q/T_1 stawiamy znak minus dla zaznaczenia, że ciepło jest oddawane przez ciało o temperaturze wyższej. Suma $-Q/T_1+Q/T_2$ jest liczbą dodatnią. Stąd wniosek, że S_2-S_1 jest nie tylko większe od zera, ale nawet większe od pewnej liczby dodatniej, czyli $S_2 > S_1$. Wyrównaniu się temperatur wspomnianych dwóch ciał towarzyszy wzrost entropii układu.

3. Obliczyć zmianę entropii ΔS, wynikającą z odwracalnego oziębienia m kg wody od T_1 do T_2. Z definicji entropii wynika

$$\Delta S = \sum_{1}^{2} \frac{\Delta Q}{T}.$$

Dzieląc całość przemiany na elementarne przejścia umożliwimy zastąpienie sumy całką

$$\Delta S = \int_{1}^{2} \frac{dQ}{T}.$$

Ale dQ możemy wyrazić jako

$$dQ = mc\,dT$$

(zmianę objętości wody podczas ogrzewania zaniedbujemy), a więc

$$\Delta S = mc \int_{T_1}^{T_2} \frac{dT}{T} = mc \ln\frac{T_2}{T_1}.$$

Ponieważ z założenia $T_2 < T_1$, więc otrzymujemy wartość ujemną o wymiarze J/K.

4. Obliczyć zmianę entropii 1 mola gazu doskonałego odpowiadającą przemianie odwracalnej od stanu początkowego (*1*), określonego parametrami (p_1, V_{1m}, T_1), do stanu końcowego (*2*) o parametrach (p_2, V_{2m}, T_2).

W tym celu korzystamy z pierwszej zasady termodynamiki:

$$dQ = dU + dW.$$

Podstawiając odpowiednie wartości zamiast dU i dW, otrzymamy

$$dQ = C_V dT + p\,dV_m.$$

Uwzględniając równanie Clapeyrona napiszemy

$$dQ = C_V dT + RT\frac{dV_m}{V_m}.$$

Jeśli nieskończenie mała odwracalna przemiana zachodzi w temperaturze T, to

$$dS = \frac{dQ}{T} = C_V\frac{dT}{T} + R\frac{dV_m}{V_m}.$$

Stąd po scałkowaniu

$$S = C_V \ln T + R \ln V_m + \text{const},$$

a zatem przejściu od stanu początkowego (*1*) do stanu końcowego (*2*) odpowiada zmiana entropii:

$$S_2-S_1 = C_V \ln\frac{T_2}{T_1} + R\ln\frac{V_{2m}}{V_{1m}}.$$

18.14. Interpretacja statystyczna zasady wzrostu entropii

Zasadę wzrostu entropii, charakteryzującą kierunek przemian samorzutnych zachodzących w układach zamkniętych, często traktuje się jako matematyczne ujęcie drugiej zasady termodynamiki.

Określoną kierunkowość przemian samorzutnych można tłumaczyć w oparciu o podstawowe założenia teorii kinetycznej budowy materii. W każdym zjawisku odbywającym się na skalę makroskopową bierze udział olbrzymia liczba cząsteczek wykonujących ruchy drgające lub poruszających się ruchem jednostajnym. Prawdopodobieństwo istnienia zupełnego braku uporządkowania tych cząsteczek tak znacznie przewyższa prawdopodobieństwo wystąpienia pewnego uporządkowania, że wszelkie samorzutne zjawiska prowadzą praktycznie zawsze do stanu większego chaosu. Wzrost entropii idzie w parze z przejściem układu do stanu mniejszego uporządkowania cząsteczek. Taką interpretację drugiej zasady termodynamiki podał Maxwell.

Stan najbardziej prawdopodobny w gazie (tzn. stan osiągany najczęściej w wyniku różnych przejść) jest to stan: 1) o równomiernym rozkładzie cząsteczek w całej objętości gazu, 2) o równomiernym rozkładzie kierunków prędkości w przestrzeni i 3) o takim rozkładzie wartości liczbowych prędkości, który odpowiada prawu Maxwella. Przejściu układu do stanu większego chaosu towarzyszy degradacja energii, gdyż układ samorzutnie nie wraca do stanu bardziej uporządkowanego, czyli mniej prawdopodobnego.

Należy jednak podkreślić, że to, co mówimy, jest słuszne z punktu widzenia statystyki. Gdyby obserwowane przez nas układy obejmowały stosunkowo małą liczbę cząsteczek (np. rzędu 1000), nie można by stosować zasady wzrostu entropii. Odnosząc obserwacje do stosunkowo małych obszarów gazowych, można stwierdzić istnienie odchyleń od II zasady termodynamiki. W dostatecznie małych elementach objętości gazu dają się zauważyć nieznaczne odchylenia od równomiernego rozkładu cząsteczek, prowadzące do powstawania wahań gęstości, zwanych *fluktuacjami gęstości*. Podobnie można mówić o *fluktuacjach temperatury*, gdy liczba cząsteczek gazu uwzględnionych przy obliczaniu średniej energii kinetycznej jest za mała.

18.15. Trzecia zasada termodynamiki

Badania właściwości ciał stałych i ciekłych w temperaturach bliskich zera bezwzględnego doprowadziły Nernsta (1906 r.) do sformułowania tzw. III zasady termodynamiki. Obecnie zwykle jest ona podawana w ujęciu Plancka. *W temperaturze zera bezwzględnego entropia ciał stałych i ciekłych staje się równa zeru*, czyli

$$\lim_{T\to 0} S = 0.$$

W oparciu o trzecią zasadę termodynamiki można teoretycznie wykazać, że w miarę zbliżania się do temperatury zera bezwzględnego ciepła właściwe ciał stałych i współczynniki rozszerzalności dążą do zera. Innymi słowy, w tych warunkach maleją różnice właściwości termicznych ciał stałych.

Pytania i zadania

1. Jakie trzy wielkości fizyczne łączy ze sobą pierwsza zasada termodynamiki?

2. Jak można scharakteryzować przemianę kołową (cykl)?

3. W jaki sposób można wykazać, że energia wewnętrzna jest funkcją stanu układu?

4. Przedstawić na jednym wykresie (V, p) cztery poznane izoprzemiany gazu doskonałego, wypisując przy każdej krzywej jej nazwę i jej charakterystykę.

5. Czemu się równa ilość energii cieplnej pobranej podczas izotermicznej przemiany w temperaturze T 1 mola gazu doskonałego od objętości V_1 do objętości V_2?

6. Jakie prawo rządzi przemianą adiabatyczną? Podać plan wyprowadzenia odpowiedniego wzoru. Wyrazić ostateczny wzór w dwóch postaciach. Przypomnieć znaczenie fizyczne współczynnika występującego w wykładniku potęgi V.

7. Jakie przemiany nazywamy przemianami politropowymi?

8. Czy zależność $dU = C_V dT$ jest spełniona dokładnie dla gazów rzeczywistych? Uzasadnić odpowiedź.

9. Zdefiniować ciepło molowe przy stałej objętości C_V i przy stałym ciśnieniu C_p. Jaki związek zachodzi między C_p, C_V i R oraz c_p, c_V i R dla określonego gazu doskonałego o masie molowej M?

10. Jakie przemiany termodynamiczne nazywamy przemianami odwracalnymi?

11. Czemu się równa maksymalna sprawność silnika pracującego odwracalnie w cyklu Carnota między temperaturami T_1 i T_2? Podać wyprowadzenie wzoru.

12. Na czym się opiera definicja termodynamicznej skali temperatur?

13. Jakie rozumowanie doprowadza do nierówności Clausiusa? Jak można ująć nierówność Clausiusa posługując się pojęciem ciepła zredukowanego?

14. Jaka jest definicja entropii? Jak można uzasadnić, że entropia jest jednoznaczną funkcją stanu układu?

15. Jakie wyrażenie przedstawia elementarną zmianę entropii dS w dowolnym odwracalnym procesie, zachodzącym w gazie doskonałym?

16. Uzasadnić twierdzenie, że proces odwracalny adiabatyczny jest procesem izoentropowym.

17. Jak można sformułować kryterium kierunkowości przemian zachodzących w przyrodzie?

18. Podać różne sformułowania drugiej zasady termodynamiki.

19. Do czego się sprowadza interpretacja statystyczna zasady wzrostu entropii?

20. Jak można sformułować trzecią zasadę termodynamiki?

21. Stalowa kulka spadająca z wysokości h_0 z prędkością początkową v_0 odbiła się po zderzeniu z podstawą na wysokość h_1. Wiedząc, że temperatura kulki wzrosła o Δt, obliczyć, jaki procent pracy zużytej na odkształcenie kulki w czasie zderzenia zamienił się na energię cieplną. Ciepło właściwe stali równa się c.

Odp. $\dfrac{200c\Delta t}{2g(h_0-h_1)+v_0^2}$.

22. Kulka ołowiana o początkowej temperaturze $t = 27°C$ topi się podczas zderzenia z niesprężystą ścianą. Czy kulka w chwili zderzenia poruszała się z prędkością większą czy mniejszą niż prędkość głosu w powietrzu w temperaturze pokojowej? Założyć, że energia kinetyczna kulki zamienia się w całości na energię cieplną, pobraną przez kulkę. Potrzebne dane znajdują się w odpowiednich tabelach.

23. Ogrzewając masę 20 g tlenu pod stałym ciśnieniem wywołano przyrost temperatury równy 100°C.

Ile na to zużyto energii cieplnej? Jaki jest przyrost energii wewnętrznej przy tym samym ogrzaniu? $C_p = 29{,}4$ J/mol · K; $C_V = 21{,}1$ J/mol · K.

Odp. 1844 J, 1315 J.

24. Dwa gazy doskonałe o temperaturze początkowej 27°C: jeden — jednoatomowy ($\varkappa_1 = 1{,}68$), drugi — dwuatomowy ($\varkappa_2 = 1{,}4$) zostały sprężone adiabatycznie do objętości równej $\frac{1}{10}$ objętości początkowej. Jakie są ich temperatury końcowe?

Odp. 1390K, 730K.

25. Ciśnienie początkowe dwuatomowego gazu $p_0 = 100$ atm. Jakie będzie ciśnienie tego gazu, gdy jego objętość się podwoi podczas rozprężania izotermicznego i adiabatycznego? Współczynnik $\varkappa = 1{,}4$.

Odp. 50 atm, 37,8 atm.

26. W zbiorniku zamkniętym ruchomym tłokiem znajduje się $m = 20$ g wodoru o temperaturze 27°C pod ciśnieniem $p = 4$ atm. Przy sprężaniu wodoru do $\frac{1}{2}$ objętości pierwotnej wykonano pracę $W = 10^5$ J i równocześnie odprowadzono do chłodnicy energię cieplną w ilości $6 \cdot 10^4$ J. Obliczyć temperaturę i ciśnienie wodoru po sprężeniu. $c_V = 10{,}15 \cdot 10^3$ J/kg · K.

Odp. 497K, 13,3 atm.

27. 1 kg powietrza o temperaturze 20°C i początkowym ciśnieniu 98 066,5 N/m^2 sprężono do ciśnienia 98 066,5 N/m^2. Ile wynosi praca sprężenia, jeżeli odbywa się ono: a) izotermicznie, b) adiabatycznie. Ciepła właściwe powietrza $c_V = 729$ J/kg · K, $c_p = 1018$ J/kg · K.

Odp. $194 \cdot 10^3$ J, $198 \cdot 10^3$ J.

28. Znaleźć zmianę entropii związaną z odwracalnym oziębieniem 2 g powietrza od 40 do 0°C a) przy stałym ciśnieniu, b) przy stałej objętości. Potrzebne dane — jak w zadaniu poprzednim.

Odp. $-0{,}278$ J/K, $-0{,}199$ J/K.

29. 2 kg wody o temperaturze 10°C odwracalnie ogrzano do 100°C i wyparowano. Obliczyć zmianę entropii.

Odp. $14{,}4 \cdot 10^3$ J/K.

CZĘŚĆ III Elektryczność i magnetyzm

Wstęp historyczny

Naukę o elektryczności i magnetyzmie rozpoczniemy od szkicu historycznego, naświetlającego rozwój tej dziedziny fizyki.

Pierwsze doświadczenia z elektryczności przeprowadzane były już w starożytności, gdyż już Tales z Miletu (600 lat p.n.e.) wspomina o tym, że potarty bursztyn wykazuje właściwości przyciągania drobnych przedmiotów. Ogólnie też znane były objawy elektryczności atmosferycznej, a mianowicie pioruny, ale natura ich była nie wyjaśniona aż do drugiej połowy XVII wieku. Wiedziano jednak, że można się ustrzec przed uderzeniem pioruna stosując wysokie, zaostrzone maszty. Podczas prac archeologicznych w Egipcie znaleziono na ścianach starożytnych świątyń napisy wyjaśniające stosowanie masztów jako środka zabezpieczającego przed „niebieskim ogniem".

Oddziaływania magnetyczne również były znane już w starożytności dzięki istnieniu w przyrodzie tzw. magnesów trwałych w postaci rudy żelaza, np. w postaci magnetytu. Magnetyt przyciągał żelazo, mógł służyć do wytwarzania sztucznych magnesów stalowych, które np. w postaci igieł magnetycznych znalazły szerokie zastosowanie jako wskaźniki kierunku północnego i południowego na Ziemi (kompasy).

Rozwój nauki o elektryczności rozpoczyna się w wieku XVIII. Wprowadzony zostaje podział ciał na przewodniki i izolatory, rozróżniona zostaje elektryczność dodatnia i ujemna, powstają maszyny elektrostatyczne, oparte na tarciu i indukcji elektrostatycznej, prowadzone są badania elektryczności atmosferycznej itp.

W roku 1785 Coulomb na podstawie doświadczeń z wagą skręceń wypowiada swoje prawa dotyczące wzajemnego oddziaływania ładunków elektrycznych i wzajemnego oddziaływania biegunów magnetycznych. Mniej więcej w tym samym czasie, dzięki odkryciu Galvaniego, zbudowane zostają pierwsze chemiczne źródła prądu elektrycznego (ogniwa Volty).

Już pierwsze dziesiątki lat XIX stulecia przynoszą bardzo ważne odkrycia, świadczące o więzi między dziedziną zjawisk elektrycznych i magnetycznych. Od tego czasu można już mówić o rozwoju elektromagnetyzmu. W roku 1820 Oersted stwierdza, że igła magnetyczna umieszczona w pobliżu przewodnika ulega odchyleniu w czasie przepływu prądu. W tym samym czasie Biot i Savart ustalają prawo określające dzia-

łanie prądu elektrycznego na pobliskie bieguny magnetyczne, a Ampère kładzie podwaliny elektrodynamiki.

Z nazwiskiem Faradaya wiąże się odkrycie prądów indukcyjnych, badanie właściwości pól elektrycznych i magnetycznych oraz badanie elektrolizy, stwarzające podstawy do nauki o elektrycznej budowie materii.

Druga połowa XIX wieku przynosi prace teoretyczne Maxwella. Jego równania stanowią podstawę klasycznej elektrodynamiki. One też były bazą do stworzenia elektromagnetycznej teorii światła. W ten sposób optyka stała się w pewnym stopniu częścią nauki o elektromagnetyzmie. Doświadczalne otrzymanie fal elektromagnetycznych w zakresie długości zwanych obecnie radiowymi udało się Hertzowi w r. 1889.

W tym samym czasie zaczynają się mnożyć liczne wynalazki wpływające na kształtowanie się warunków życia ludzkiego. Odkrycie prądów indukcyjnych umożliwiło budowę maszyn elektrycznych, tj. urządzeń przetwarzających energię elektryczną na energię mechaniczną i odwrotnie. Maszyny takie zrewolucjonizowały technikę napędu elektrycznego przyczyniając się w ten sposób decydująco do szybkiego rozwoju różnych dziedzin przemysłu.

Do otrzymywania energii elektrycznej udało się wykorzystać energię cieplną, energię wód (hydroelektrownie) i energię chemiczną. Ze względu na to, że energię elektryczną można przesyłać ze stosunkowo małymi stratami na duże odległości, znajduje ona coraz szersze zastosowanie. Trudno dziś wyliczyć wszystkie dziedziny życia, w których korzystamy z energii prądu elektrycznego. Ogólnie znany jest rozwój techniki oświetleniowej od lamp łukowych węglowych, przez lampy żarzenia do nowoczesnych lamp jarzeniowych. Przykładowo można też wspomnieć o zastosowaniu energii elektrycznej w telegrafii, telefonii, radiotechnice, telewizji, elektroautomatyce itp.

Rozwinięcie elektrycznych metod badawczych pozwoliło człowiekowi pogłębić badania eksperymentalne w różnych dziedzinach i wniknąć głębiej w budowę materii. Wymienić tu można odkrycie elektronu, promieni kanalikowych, promieni Röntgena, odkrycie zjawisk termo- i fotoemisji, promieniotwórczości naturalnej i sztucznej, dezintegracji ciężkich jąder, odkrycie neutronu, pozytonu, promieni kosmicznych, mezonów, hiperonów, antycząstek itp. Rozwijają się nowe, szerokie dziedziny badań i zastosowań, jak elektronika, fizyka ciała stałego itp.

Równocześnie z rozwojem badań eksperymentalnych odbywa się rozwój teorii. Okazało się, że wielu zjawisk nie można wyjaśnić w oparciu o dawne teorie klasyczne. Konieczne staje się wprowadzenie pojęcia kwantów i stworzenie mechaniki falowej, umożliwiającej matematyczne ujęcie zjawisk fizycznych.

Przedstawienie tematyki z dziedziny elektromagnetyzmu w podręczniku rozpoczynamy od elektrostatyki, następnie mówimy o prądzie elektrycznym stałym, polu magnetycznym, zjawisku indukcji elektromagnetycznej i prądzie elektrycznym zmiennym i na zakończenie — o drganiach i falach elektromagnetycznych.

Wysunięcie na plan pierwszy elektrostatyki skupi naszą uwagę na ładunku elektrycznym i wytwarzanym przez niego polu elektrostatycznym. Jak wiemy jednak z § 2 i 3 „Wiadomości wstępnych", w obowiązującym obecnie międzynarodowym układzie jednostek SI podstawową wielkością elektryczną jest nie ładunek, lecz natężenie prądu

elektrycznego, zwane krótko prądem elektrycznym. Podstawową jednostką elektryczną jest 1 amper (por. definicję ampera podaną w § 3 „Wiadomości wstępnych"). Na podstawie definicji prądu elektrycznego $I = dQ/dt$ można ustalić związek między jednostką ładunku w układzie SI (kulombem) i jednostką prądu (amperem)

$$1\,\mathrm{C} = 1\,\mathrm{A} \cdot 1\,\mathrm{s}.$$

Niektóre z wzorów elektrostatycznych przedstawionych w tym rozdziale będą się nieco różniły od wzorów znanych ze szkoły średniej, właśnie z racji dostosowania ich do obowiązującego układu SI.

ROZDZIAŁ 19

Elektrostatyka

19.1. Ładunek elektryczny

Przegląd wielkości, praw i zjawisk elektrycznych rozpoczniemy od ładunku elektrycznego. Jak wiadomo, rozróżniamy dwa rodzaje ładunków: ładunki dodatnie i ujemne. Przypisanie znaków + i — ładunkom elektrycznym było zupełnie dowolne. Umówiono się, że szkło pocierane jedwabiem ładuje się dodatnio, a ebonit pocierany suknem — ujemnie. Ładunków nie potrafimy wytwarzać: umiemy tylko rozdzielać ładunki dodatnie i ujemne, występujące w przyrodzie. Z niemożliwością wytwarzania ładunków ściśle się wiąże jedno z podstawowych praw przyrody, należące do grupy tzw. *zasad zachowania*, a mianowicie *zasada zachowania ładunku**. Tak np. przy pocieraniu szkła jedwabiem szkło ładuje się dodatnio, a jedwab — ujemnie, przy czym ładunek dodatni szkła i ujemny jedwabiu są sobie równe.

Jako interesujący przykład zasady zachowania ładunku można wysunąć tzw. *proces anihilacji*. Jest to reakcja zachodząca między dwiema cząstkami elementarnymi: elektronem o ładunku $-e$ i pozytonem o ładunku $+e$, w wyniku której obie cząstki znikają, a na ich miejsce pojawiają się dwa fotony promieniowania γ. Ładunek wypadkowy obu cząstek przed i po anihilacji wynosi zero: spełniona jest zatem zasada zachowania ładunku. Ponadto w czasie procesu anihilacji łączna masa obu cząstek ulega zamianie na energię promieniowania zgodnie ze znanym już nam wzorem $E = mc^2$.

Do zasady zachowania ładunku będziemy się stale odwoływali przy pisaniu równań rozpadu promieniotwórczego i reakcji jądrowych.

Ładunki dodatnie i ujemne istnieją w każdym atomie. *Atom* bowiem składa się z dodatnio naładowanego jądra i z krążących dookoła niego ujemnie naładowanych elektronów. Gdy atom jest w stanie obojętnym, liczba elektronów jest taka, że ich łączny ładunek ujemny neutralizuje całkowity dodatni ładunek jądra. Gdy atom traci jeden lub więcej elektronów, powstaje *jon* dodatni jedno- lub kilkuwartościowy. Przyłączenie do atomu nadprogramowych elektronów prowadzi do wytworzenia jonów ujemnych.

* Przypominamy, że już poprzednio poznaliśmy *zasadę zachowania energii i masy*, *zasadę zachowania pędu* i *zasadę zachowania momentu pędu*.

Ciała występujące w przyrodzie dzielimy pod względem elektrycznym na trzy grupy: *przewodniki*, *dielektryki* (izolatory) i *półprzewodniki*. Do grupy pierwszej zaliczamy m. in. metale, grafit, elektrolity. W przewodnikach metalicznych istnieją elektrony swobodne decydujące o dobrym przewodnictwie elektrycznym. Tylko wtedy, gdy przewodnik metaliczny jest izolowany od ziemi, można go trwale naelektryzować przez zetknięcie z innym naelektryzowanym przewodnikiem lub przez indukcję. W przeciwnym przypadku swobodne elektrony spływają z przewodnika do ziemi.

Do grupy dielektryków należą: bursztyn, kwarc, porcelana, ebonit, jedwab, masy plastyczne itd. W dielektrykach nie ma swobodnych elektronów. Przez pocieranie dowolnego dielektryka innym wywołujemy trwałe naelektryzowanie zarówno ciała pocieranego, jak i ciała pocierającego. Na obu ciałach występują ładunki równe, ale przeciwnego znaku.

Do grupy półprzewodników zaliczamy m. in. krzem, german, selen, liczne tlenki metali, azotki, węgliki. Właściwości tych ciał są w dużym stopniu zależne od sposobu ich otrzymywania, od naświetlania, od istniejących w nich choćby minimalnych domieszek obcych lub domieszek wynikających z niezachowania stosunków stechiometrycznych. Opór właściwy tych ciał w dużym stopniu zależy od temperatury, przy czym w przeciwieństwie do metali opór półprzewodników na ogół maleje ze wzrostem temperatury.

Wyżej wymienione grupy ciał można scharakteryzować przez podanie granicznych wartości oporu właściwego lub przewodności właściwej. Przewodność właściwa metali zmienia się w granicach 10^8–10^6 $\Omega^{-1}\cdot m^{-1}$, przewodność właściwa izolatorów w granicach 10^{-9}–10^{-18} $\Omega^{-1}\cdot m^{-1}$, a przewodność właściwa półprzewodników w granicach 10^7–10^{-8} $\Omega^{-1}\cdot m^{-1}$. Bliższą charakterystykę metali, półprzewodników i izolatorów podamy w § 32.8.

Doświadczenie uczy, że ciała naelektryzowane jednoimiennie wzajemnie się odpychają, a naelektryzowane różnoimiennie wzajemnie się przyciągają. Siłę wzajemnego oddziaływania ładunków Q_1 i Q_2 znajdujących się w odległości r określił Coulomb opierając się na doświadczeniach z wagą skręceń. Budowa tego przyrządu znana jest z kursu szkoły średniej. *Prawo Coulomba*, dotyczące w zasadzie ładunków punktowych, można zapisać w następującej postaci:

$$F \sim \frac{Q_1 Q_2}{r^2} \quad \text{lub} \quad F = k\frac{Q_1 Q_2}{r^2}, \tag{19.1}$$

gdzie k oznacza współczynnik proporcjonalności zależny od doboru jednostek i od właściwości ośrodka, w którym przeprowadza się badania.

Gdybyśmy we wzorze (19.1) mieli swobodę wyboru którejkolwiek z jednostek, można by przyjąć taką definicję, by współczynnik k równał się 1. W układzie SI mamy jednak już narzucone: jednostkę siły — niuton, jednostkę odległości — metr i jednostkę ładunku — kulomb. A zatem współczynnik k musi mieć taką wartość, aby równanie (19.1) było spełnione dla wszelkich wyników doświadczalnych. Biorąc np. pod uwagę dwa jednakowe, punktowe ładunki równe 1 kulombowi, rozsunięte na odległość 1 metra w próżni, można doświadczalnie wyznaczyć siłę ich wzajemnego oddziaływania F

w niutonach, następnie podstawić dane do wzoru (19.1) i obliczyć współczynnik k_0 dla próżni. Otrzymuje się w ten sposób wartość:

$$k_0 = 9{,}0 \cdot 10^9\, \mathrm{N \cdot m^2 \cdot C^{-2}}. \tag{19.2}$$

W dawniej stosowanych układach jednostek w wielu wzorach z dziedziny elektryczności występował czynnik 4π nie uzasadniony np. względami symetrii przestrzennej. Okazało się, że przez wprowadzenie tzw. *racjonalizacji* można spowodować zniknięcie czynnika 4π z równań, w których jego obecność nie jest uzasadniona, a pojawienie się jego w takich równaniach, gdzie ta obecność jest uzasadniona. Racjonalizacja w odniesieniu do wzoru Coulomba wiąże się z nadaniem współczynnikowi k ogólnej postaci

$$k = \frac{1}{4\pi\varepsilon} \tag{19.3}$$

lub w przypadku próżni

$$k_0 = \frac{1}{4\pi\varepsilon_0}, \tag{19.4}$$

gdzie ε i ε_0 oznaczają odpowiednio *przenikalność elektryczną danego ośrodka i przenikalność elektryczną próżni* (przenikalność elektryczna jest także nazywana *stałą dielektryczną*).

Korzystając z (19.2) i (19.4) łatwo znajdziemy ε_0:

$$\varepsilon_0 = \frac{1}{4\pi k_0} = 8{,}85415 \cdot 10^{-12} \mathrm{C^2 \cdot N^{-1} \cdot m^{-2}} \approx 8{,}9 \cdot 10^{-12} \mathrm{C^2 \cdot N^{-1} \cdot m^{-2}}. \tag{19.5}$$

Wzór Coulomba zracjonalizowany w odniesieniu do dowolnego ośrodka ma zatem postać

$$F = \frac{1}{4\pi\varepsilon}\frac{Q_1 Q_2}{r^2}, \tag{19.6}$$

a w odniesieniu do próżni —

$$F = \frac{1}{4\pi\varepsilon_0}\frac{Q_1 Q_2}{r^2}. \tag{19.7}$$

Czasem wygodnie jest posługiwać się pojęciem *względnej przenikalności elektrycznej ε_r danego ośrodka w stosunku do próżni*:

$$\varepsilon_r = \frac{\varepsilon}{\varepsilon_0}. \tag{19.8}$$

Przenikalność elektryczną (*bezwzględną*) ε danego ośrodka można zatem wyrazić jako

$$\varepsilon = \varepsilon_0 \varepsilon_r. \tag{19.9}$$

Wzór (19.6) można więc przedstawić w postaci

$$F = \frac{1}{4\pi\varepsilon_0 \varepsilon_r}\frac{Q_1 Q_2}{r^2}. \tag{19.10}$$

Z porównania wzorów (19.7) i (19.10) widać, że względna przenikalność elektryczna danego ośrodka jest to liczba wskazująca, ile razy zmniejsza się siła wzajemnego oddziaływania dwóch ładunków, gdy nie zmieniając ich odległości przeniesiemy je z próżni do danego ośrodka. Wartości liczbowe względnych przenikalności elektrycznych podano w tab. 19.2 w § 19.16, gdzie też *szczegółowiej omówiono właściwości dielektryków.* Tutaj tylko jeszcze podkreślimy, że względna przenikalność elektryczna powietrza $\varepsilon_r = 1{,}000\,59 \approx 1$, a więc wzory odniesione do próżni są z dużym stopniem dokładności słuszne i dla powietrza.

Tak więc *wskutek wprowadzenia racjonalizacji i uwzględnienia obowiązującego układu* SI *we wzorze Coulomba pojawił się czynnik* 4π *i przenikalność elektryczna próżni została zdefiniowana jako liczba mianowana, nierówna jednostce, mająca wartość liczbową wymienioną we wzorze* (19.5).

Siła wzajemnego oddziaływania dwóch ładunków jest siłą centralną skierowaną wzdłuż prostej łączącej te ładunki. Ze wzoru Coulomba wynika, że przyciąganiu ładunków różnoimiennych odpowiada siła ujemna, a odpychaniu ładunków jednoimiennych — siła dodatnia.

Jeżeli w otoczeniu ładunku Q_1 znajduje się kilka innych ładunków Q_2, Q_3, Q_4 ..., to każdy z nich działa na Q_1 siłą wynikającą ze wzoru Coulomba, niezależnie od liczby i rozmieszczenia ładunków pozostałych. Wypadkowa siła $\mathbf{F}_1$ działająca na ładunek Q_1 jest sumą geometryczną wymienionych sił składowych:

$$\mathbf{F}_1 = \mathbf{F}_2+\mathbf{F}_3+\mathbf{F}_4+ \ldots +\mathbf{F}_n.$$

Mówiąc o ładunkach Q_1, Q_2, ... należy pamiętać o bardzo istotnej właściwości ładunku, a mianowicie o jego kwantowości, sprowadzającej się do istnienia pewnej najmniejszej, dalej niepodzielnej porcji ładunku. Jest to tzw. *ładunek elementarny* $e = 1{,}602\,06 \cdot 10^{-19}$ C, który występuje w przyrodzie, np. jako ujemny ładunek elektronu, dodatni ładunek protonu lub pozytonu, dodatni lub ujemny ładunek jednowartościowego jonu. Sposób wyznaczania ładunku elementarnego opisano w § 19.17. Wszystkie ładunki występujące w przyrodzie są tylko całkowitymi wielokrotnościami ładunku elementarnego.

Tak na przykład atom pierwiastka o liczbie porządkowej Z składa się z jądra zawierającego Z protonów, a więc mającego ładunek Ze kulombów. Ładunek jądra jest zrównoważony ładunkiem $-Ze$, który należy przypisać Z elektronom wchodzącym w skład atomu. Siłami wiążącymi w atomie są siły elektryczne, działające między jądrem i elektronami.

Może interesujące będzie porównanie sił elektrycznych i grawitacyjnych, występujących w atomach. Na przykład w atomie wodoru złożonym z 1 protonu ($m_p \doteq 1{,}672 \cdot 10^{-27}$ kg) i 1 elektronu ($m_e = 9{,}108 \cdot 10^{-31}$ kg) wzajemna odległość tych cząstek szacowana jest na $5{,}3 \cdot 10^{-11}$ m. Łatwo sprawdzić stosując prawo Coulomba i prawo powszechnego ciążenia, że siła przyciągania elektrycznego $F_e = 8{,}1 \cdot 10^{-8}$ N, a siła grawitacyjna $F_g = 3{,}7 \cdot 10^{-47}$ N. Innymi słowy, siła grawitacyjna jest w tym przypadku około 10^{39} razy mniejsza.

19.2. Pole elektrostatyczne. Natężenie pola

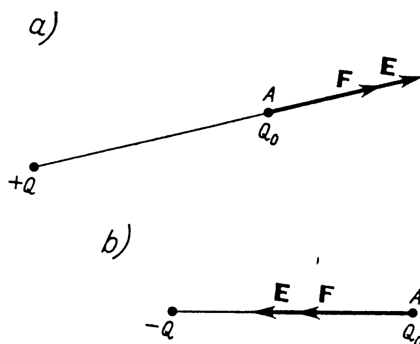
O istocie pola elektrostatycznego mówiliśmy już w § 4.6. Wiemy, że ładunek elektryczny wywołuje w swoim otoczeniu pewne zmiany objawiające się w ten sposób, że inny ładunek wprowadzony do tego obszaru podlega działaniu sił kulombowskich. Mówimy, że ładunek elektryczny wytwarza dookoła siebie tzw. pole elektryczne. Do badania pola służą ładunki próbne tak małe, by swoją obecnością nie zakłócały rozkładu ładunków pozostałych. Jeśli siła działająca na ładunek próbny w danym punkcie pola nie ulega zmianie z biegiem czasu, mówimy o elektrycznym polu stacjonarnym lub — krócej — o polu elektrostatycznym.

Pole elektryczne można — podobnie jak pole grawitacyjne — opisywać wektorowo i skalarnie. Przedstawienie wektorowe pola sprowadza się do wyznaczania natężenia pola, przedstawienie skalarne zaś do wyznaczania potencjału w poszczególnych punktach pola. Na razie zajmiemy się przedstawieniem wektorowym.

Przez *natężenie pola elektrycznego* **E** w danym punkcie rozumiemy stosunek siły działającej w tym punkcie na ładunek próbny dodatni Q_0 do wartości tego ładunku. A zatem

$$\mathbf{E} = \frac{\mathbf{F}}{Q_0}. \tag{19.11}$$

Jest to równanie definicyjne o charakterze wektorowym: wektor natężenia pola **E** ma kierunek zgodny z kierunkiem siły **F**, działającej na dodatni ładunek próbny Q_0 (rys. 19.1).



Rys. 19.1

Natężenie pola równa się jednostce (w układzie SI) w takim punkcie, w którym na ładunek jednego kulomba działa siła jednego niutona. Jednostka ta nie ma specjalnej nazwy: natężenie pola wyrażamy w N/C*.

Z równania (19.11) wynika zależność

$$\mathbf{F} = Q_0\mathbf{E}. \tag{19.12}$$

* Trzeba od razu podkreślić, że najczęściej stosowaną nazwą tej jednostki natężenia pola elektrycznego jest wolt na metr (V/m). Do tej sprawy jeszcze wrócimy w § 19.10.

Jeżeli badane pole jest wytwarzane przez ładunek punktowy Q, to uwzględniając prawo Coulomba można napisać

$$E = \frac{1}{4\pi\varepsilon}\frac{QQ_0}{r^2 Q_0} = \frac{1}{4\pi\varepsilon}\frac{Q}{r^2}. \tag{19.13}$$

Jak widać, natężenie pola wytwarzanego przez ładunek Q zależy nie tylko od Q i r, lecz także od właściwości ośrodka za pośrednictwem jego przenikalności elektrycznej ε.

W przypadku gdy pole elektryczne wytwarzane jest przez kilka ładunków punktowych $Q_1, Q_2, Q_3, \ldots, Q_n$, natężenie pola w dowolnym punkcie obliczamy sumując wektorowo natężenia $\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3, \ldots, \mathbf{E}_n$, pochodzące od poszczególnych ładunków i obliczone przy założeniu, że każdy z nich jest jedynym ładunkiem wytwarzającym pole:

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 + \mathbf{E}_3 \ldots + \mathbf{E}_n.$$

Przykład. Obliczyć natężenie pola w otoczeniu tzw. *dipola elektrycznego*, tj. układu dwóch ładunków różnoimiennych, jednakowych co do wartości $+Q$ i $-Q$, rozsuniętych na odległość a, biorąc pod uwagę: 1) punkty leżące na osi dipola (rys. 19.2), 2) punkty leżące na symetralnej dipola (rys. 19.3).

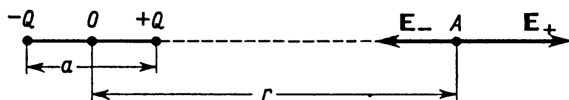
1. Weźmy pod uwagę punkt A leżący na osi dipola w odległości r od jego środka O. Natężenie $\mathbf{E}_A$ pola w punkcie A jest wypadkową natężeń pól wytwarzanych w punkcie A przez ładunek $+Q$ i $-Q$. Oba te natężenia $\mathbf{E}_+$ i $\mathbf{E}_-$ są skierowane wzdłuż tej samej prostej, lecz mają zwroty przeciwne, a zatem ich suma geometryczna sprowadza się do różnicy arytmetycznej:

$$E_A = \frac{1}{4\pi\varepsilon}\frac{Q}{(r-a/2)^2} - \frac{1}{4\pi\varepsilon}\frac{Q}{(r+a/2)^2},$$

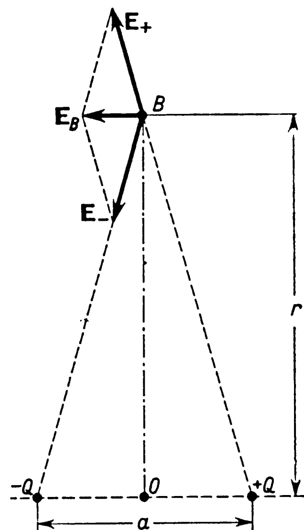
$$E_A = \frac{Q}{4\pi\varepsilon}\frac{(r+a/2)^2-(r-a/2)^2}{(r+a/2)^2(r-a/2)^2},$$

$$E_A = \frac{Q}{4\pi\varepsilon}\frac{r^2+ra+a^2/4-r^2+ra-a^2/4}{(r^2-a^2/4)^2},$$

$$E_A = \frac{1}{4\pi\varepsilon}\frac{2Qra}{(r^2-a^2/4)^2}.$$



Rys. 19.2



Rys. 19.3

Takie jest wyrażenie ogólne na natężenie pola w punktach leżących na osi dipola. Dla punktów leżących daleko od ładunków dipola (tzn. gdy $r \gg a$) otrzymujemy wzór przybliżony

$$E \approx \frac{1}{4\pi\varepsilon}\frac{2Qa}{r^3}.$$

Iloczyn ładunku Q dipola i odległości a nazywamy *momentem dipola*. Tę nową wielkość traktujemy jako wektor o kierunku od ładunku ujemnego do ładunku dodatniego dipola i oznaczamy symbolem $\mathbf{p}$. A zatem natężenie pola elektrycznego w punktach leżących na osi dipola w odległości r znacznie większej od

odległości a wyraża się przybliżonym wzorem

$$\mathbf{E} = \frac{1}{4\pi\varepsilon}\frac{2\mathbf{p}}{r^3}.$$

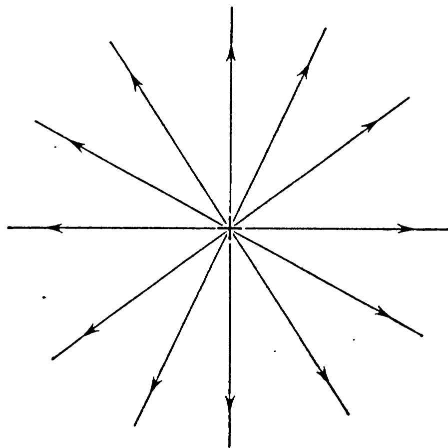
2. W przypadku drugim natężenie pola $\mathbf{E}_B$ w punkcie B leżącym na symetralnej dipola znajdziemy jako sumę wektorową natężeń $\mathbf{E}_+$ i $\mathbf{E}_-$ (rys. 19.3). Proponujemy samodzielne sprawdzenie, że przy założeniu $r \gg a$ obowiązuje wzór przybliżony

$$E = \frac{1}{4\pi\varepsilon}\frac{p}{r^3}.$$

19.3. Linie sił pola elektrycznego. Strumień elektryczny

Do modelowego opisu pola elektrycznego stosuje się jeszcze obecnie wprowadzone przez Faradaya tzw. linie sił pola elektrycznego, które dalej będziemy krótko nazywali liniami sił. *Linia siły* jest to linia o takim przebiegu, że styczna do niej poprowadzona w dowolnym jej punkcie wyznacza kierunek wektora natężenia pola elektrycznego w tym punkcie. Ujmując zagadnienie bardziej obrazowo można powiedzieć, że linia siły przedstawia tor ruchu swobodnego, pozbawionego masy, ładunku próbnego elektrycznego dodatniego, wprowadzonego do pola. Kierunek zaznaczany strzałką na linii siły to właśnie kierunek ruchu wymienionego ładunku.

Układ linii sił w kilku prostych przypadkach przedstawiają rys. 19.4–19.7.

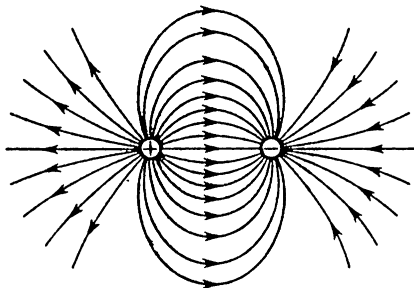


Rys. 19.4

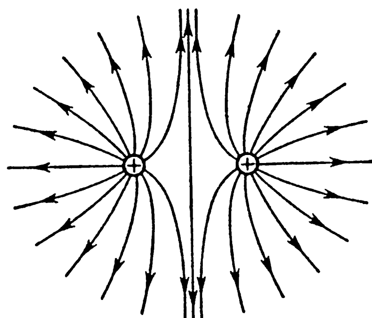
Rysunek 19.4 przedstawia radialny rozbieżny układ linii sił w otoczeniu pojedynczego ładunku punktowego dodatniego: wszystkie linie zaczynają się na tym ładunku i rozbiegają promieniście w przestrzeń. Układ linii w otoczeniu pojedynczego ładunku punktowego ujemnego byłby radialny zbieżny: wszystkie linie z przestrzeni dochodziłyby do ładunku ujemnego i na nim się kończyły.

Rysunek 19.5 przedstawia układ linii sił w otoczeniu dipola elektrycznego: wszystkie linie zaczynają się na ładunku dodatnim, a kończą na ujemnym.

Rysunek 19.6 dotyczy układu linii sił w otoczeniu dwóch ładunków dodatnich jednakowych co do wartości. Linie sił z obu ładunków wybiegają w przestrzeń nigdzie się nie przecinając. Gdyby oba ładunki były ujemne, rozkład przestrzenny linii sił pozostałby niezmieniony, zmianie uległyby tylko kierunki linii.



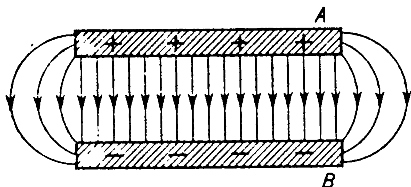
Rys. 19.5



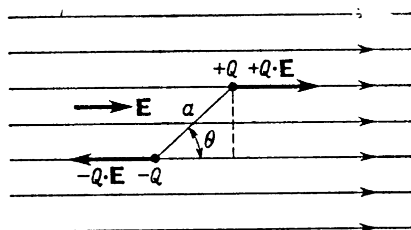
Rys. 19.6

Rysunek 19.7 przedstawia układ linii sił w otoczeniu kondensatora płaskiego. Między okładkami A i B kondensatora linie biegną równolegle do siebie. Taki przebieg jest charakterystyczny dla tzw. ***pola jednorodnego***, tzn. takiego, w którym natężenie pola **E** jest jednakowe w każdym punkcie. Jak widać z rysunku, jednorodność pola jest zakłócona w pobliżu krawędzi okładek kondensatora płaskiego.

Przebieg linii sił pola elektrycznego można badać doświadczalnie stos[illegible] ıałe swobodne dipole. Zacznijmy od przeanalizowania zachowania się pojedyncz[illegible] swobodnego dipola w jednorodnym polu elektrycznym (rys. 19.8).



Rys. 19.7



Rys. 19.8

Niech oś dipola tworzy początkowo pewien kąt θ z kierunkiem linii sił. Na oba ładunki dipola działają siły równe co do wartości, lecz przeciwne co do znaku (por. wzór (19.12)). Siły te tworzą parę sił, której moment M równa się:

$$M = QEa\sin\theta.$$

Ale, jak już wiemy, iloczyn $Qa = p$ (moment dipola), a zatem

$$M = pE\sin\theta$$

albo w postaci wektorowej

$$\mathbf{M} = \mathbf{p}\times\mathbf{E}.$$

Pod działaniem tego momentu skręcającego swobodny dipol obraca się do położenia zgodnego z kierunkiem linii siły. W takim położeniu moment siły równa się zeru, a obie siły działające usiłują tylko szerzej rozsunąć ładunki dipola.

Od dipola pojedynczego przejdźmy do zbioru dipoli o bardzo małych odległościach a. Przy takim założeniu można przyjmować, że każdy z dipoli znajduje się w obszarze jednorodnego pola elektrycznego, a zatem każdy z osobna zachowuje się tak, jak poprzednio opisaliśmy. Dipole zbioru ułożą się wzdłuż linii sił pola elektrycznego. Wyobraźmy sobie np., że taki zbiór małych dipoli stanowi zawiesina drobnych nasion w izolującej cieczy, w której pole elektryczne jest wytworzone przez ładunek punktowy. Układ dipoli odpowiada w tym przypadku schematycznemu rysunkowi 19.4.

W zasadzie linie sił pola elektrycznego można by prowadzić dowolnie gęsto. Podobnie jak w odniesieniu do linii prądu w polu przepływu płynu (por. § 9.5), tak i w polu elektrostatycznym wprowadzono umowę normującą gęstość linii. Umowa dotyczy tzw. gęstości strumienia linii. *Przez strumień ψ_E linii sil pola elektrycznego* (zwany też krótko *strumieniem elektrycznym*) przez daną powierzchnię rozumiemy liczbę linii sił przebijających tę powierzchnię. Przez *gęstość strumienia elektrycznego* rozumiemy stosunek $d\psi_E/dS_\perp$, przy czym $dS_\perp$ jest elementarną powierzchnią ustawioną prostopadle do kierunku linii pola. Według umowy gęstość strumienia elektrycznego jest równa wartości liczbowej natężenia pola elektrycznego na tej powierzchni. A zatem

$$\frac{d\psi_E}{dS_\perp} = E.$$

Stąd elementarny strumień elektryczny przez powierzchnię $dS_\perp$ równa się

$$d\psi_E = E dS_\perp. \tag{19.14}$$

Jeśli badana powierzchnia dS ustawiona jest ukośnie względem linii sił, tak że normalna do niej tworzy kąt α z wektorem $\mathbf{E}$, bierzemy pod uwagę tylko jej rzut $dS_\perp$ na płaszczyznę prostopadłą do $\mathbf{E}$, czyli

$$dS_\perp = dS \cos\alpha.$$

W tym ogólnym przypadku wyrażenie (19.14) sprowadza się do postaci

$$d\psi_E = E dS \cos\alpha. \tag{19.15a}$$

Jeśli umówimy się traktować $d\mathbf{S}$ jako wektor o module $|d\mathbf{S}|$ i o kierunku normalnej do powierzchni, to powyższą zależność można napisać w postaci wektorowej

$$d\psi_E = \mathbf{E} \cdot d\mathbf{S}, \tag{19.15b}$$

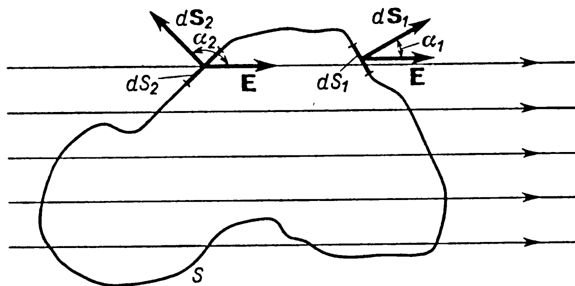
gdzie iloczyn $\mathbf{E} \cdot d\mathbf{S}$ jest iloczynem skalarnym wektorów $\mathbf{E}$ i $d\mathbf{S}$. Wielkość $d\psi_E$ jest oczywiście wielkością skalarną. Do zagadnienia zwrotu wektora $d\mathbf{S}$ wrócimy niżej.

Obliczenie całkowitego strumienia elektrycznego ψ_E przez skończoną powierzchnię S sprowadza się do wektorowej całki powierzchniowej rozciągniętej na powierzchnię S:

$$\psi_E = \int \mathbf{E} \cdot d\mathbf{S}. \tag{19.16}$$

Strumień elektryczny wyraża się w $\mathrm{N \cdot m^2/C}$ (w układzie SI).

Wróćmy jeszcze do zwrotu przypisywanego wektorowi $d\mathbf{S}$. Ustalenie tego zwrotu jest ważne, gdy całkowita powierzchnia S, względem której obliczany jest strumień ψ_E, stanowi powierzchnię zamkniętą. Wtedy bowiem rozróżniać można linie „wychodzące" z wnętrza powierzchni i „wchodzące" do niej z zewnątrz. Umówiono się, że wkład do ogólnej wartości strumienia elektrycznego wnoszony przez linie sił wychodzące z wnętrza powierzchni będzie traktowany jako dodatni. Z wzoru (19.15) wynika, że ten wkład na elementarnej powierzchni dS_1 (rys. 19.9) będzie dodatni, gdy kąt między $\mathbf{E}$ i normalną do powierzchni dS_1 będzie ostry. Jeśli — jak na rys. 19.9 — linie sił przebijają powierzchnię dS_1 i wychodzą na zewnątrz, to zwrot $\mathbf{E}$ jest od powierzchni na zewnątrz i taki sam jest zwrot $d\mathbf{S}_1$. Na powierzchni dS_2 mamy sytuację odmienną: wchodzące z zewnątrz linie sił decydują o tym, że elementarny strumień elektryczny jest na tej powierzchni ujemny, a zatem kąt między $\mathbf{E}$ i $d\mathbf{S}_2$ musi być rozwarty. Ale tym razem zwrot $\mathbf{E}$ jest do wnętrza zamkniętej powierzchni S. A zatem i w tym przypadku wektor $d\mathbf{S}_2$ ma zwrot zgodny z normalną zewnętrzną. Innymi słowy, zwrot wektora $d\mathbf{S}$ jest zawsze zgodny z normalną zewnętrzną do danej powierzchni, niezależnie od tego, czy przez tę powierzchnię elementarną linie wchodzą do wnętrza, czy też wychodzą na zewnątrz zamkniętej powierzchni S.



Rys. 19.9

19.4. Prawo Ostrogradskiego–Gaussa

Prawo Ostrogradskiego–Gaussa, zwane też krótko *prawem Gaussa*, dotyczy zależności strumienia elektrycznego przechodzącego przez dowolną zamkniętą powierzchnię od ogólnego ładunku Q znajdującego się w obszarze objętym tą powierzchnią. Dowód prawa Gaussa podamy dla powierzchni kulistej o promieniu R (rys. 19.10), w której środku znajduje się ładunek $+Q$. Linie sił wychodzą radialnie z tego ładunku i przecinają prostopadle powierzchnię kuli. Natężenie pola E w dowolnym punkcie tej powierzchni równa się

$$E = \frac{1}{4\pi\varepsilon}\frac{Q}{R^2}.$$

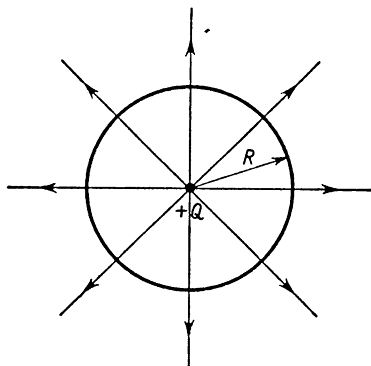
Strumień elektryczny przez powierzchnię kuli wynosi zatem

$$\psi_E = \oint \mathbf{E}\cdot d\mathbf{S} = \frac{1}{4\pi\varepsilon}\frac{Q}{R^2}4\pi R^2,$$

czyli

$$\psi_E = \frac{Q}{\varepsilon}. \tag{19.17}$$

Jak widać z otrzymanego wzoru, całkowity strumień elektryczny nie zależy od promienia kuli, przez którą przechodzi, a zależy od ładunku Q znajdującego się wewnątrz i od przenikalności elektrycznej ośrodka. Można udowodnić, że wzór Gaussa (19.17) nie zmienia swej postaci przy zastąpieniu kuli dowolną zamkniętą powierzchnią.



Rys. 19.10

Jeżeli wewnątrz zamkniętej powierzchni znajduje się n ładunków $Q_1, Q_2, Q_3, \ldots, Q_n$ (dodatnich i ujemnych), to całkowity strumień elektryczny przechodzący przez tę powierzchnię wynosi

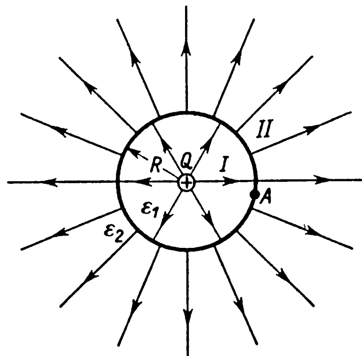
$$\psi_E = \frac{1}{\varepsilon}(Q_1+Q_2+Q_3+ \ldots +Q_n) = \frac{1}{\varepsilon}\sum_{i=1}^{n} Q_i. \tag{19.18}$$

Jeżeli powierzchnia zamknięta obejmuje ładunki dodatnie i ujemne w takiej ilości, że ich suma algebraiczna równa się zeru, to całkowity strumień elektryczny przez tę powierzchnię też równa się zeru. Wynika to z tego, że strumieniom cząstkowym przypisaliśmy (por. § 19.3) znaki + i —. Linie sił, zaczynające się na ładunkach dodatnich, przebijające powierzchnię i wychodzące na zewnątrz, dają dodatni wkład do całkowitego strumienia, natomiast linie wchodzące z zewnątrz — wkład ujemny. Może się zdarzyć, że badana powierzchnia nie obejmuje żadnych ładunków, ale jest przecinana przez linie sił. Tym razem każda z linii brana jest pod uwagę dwukrotnie: raz — jako wnosząca wkład dodatni, drugi raz — ujemny. A zatem całkowity strumień w tym przypadku równa się zeru.

Raz jeszcze podkreślamy, że strumień elektryczny przechodzący przez dowolną powierzchnię zamkniętą zależy od sumy algebraicznej ładunków znajdujących się w obszarze tej powierzchni (niezależnie od ich rozmieszczenia) i od przenikalności elektrycznej ε ośrodka, w którym pomyślana jest dana powierzchnia. Przypisanie całemu ośrodkowi przenikalności elektrycznej równej ε jest równoznaczne z założeniem jednorodności i nieograniczoności danego ośrodka.

19.5. Natężenie pola elektrycznego na granicy dielektryków

Opis pola elektrycznego za pomocą linii sił bardzo się komplikuje, gdy w rozważanym obszarze znajduje się granica dwóch ośrodków o przenikalnościach elektrycznych ε_1 i ε_2. Jak za chwilę wykażemy, *na granicy dwóch ośrodków występuje nieciągłość przebiegu linii sił*. W tych warunkach bardziej przydatny do opisu pola elektrycznego jest wektor indukcji elektrostatycznej **D**, zwany też wektorem przesunięcia elektrycznego. Zanim przejdziemy do zdefiniowania tego wektora zajmiemy się nieciągłością linii sił na granicy dwóch ośrodków.



Rys. 19.11

Rozważmy nieciągłość linii sił na następującym przykładzie. Niech ośrodek *I* w postaci ciała kulistego (np. kropli cieczy) o promieniu R (rys. 19.11) i przenikalności elektrycznej ε_1 graniczy z ośrodkiem *II* o przenikalności elektrycznej ε_2. Ładunek $+Q$ niech znajduje się w środku kuli. Weźmy pod uwagę dowolny punkt A, leżący na powierzchni rozdzielającej oba ośrodki. Zakładając, że punkt ten należy do ośrodka *I*, natężenie pola E'_A wyrazimy wzorem

$$E'_A = \frac{1}{4\pi\varepsilon_1}\frac{Q}{R^2},$$

natomiast przyjmując punkt A za punkt należący do ośrodka *II* znajdziemy

$$E''_A = \frac{1}{4\pi\varepsilon_2}\frac{Q}{R^2}.$$

Jak z tego widać, przy przekraczaniu powierzchni granicznej mamy skok wartości **E**. Ale — zgodnie z przyjętą umową (por. § 19.3) — wartości E przedstawiają gęstości strumienia elektrycznego, czyli gęstości linii sił. A zatem na powierzchni granicznej występuje rzeczywiście nieciągłość linii sił pola elektrycznego.

Wróćmy jeszcze do równań wyrażających E'_A i E''_A. Z przekształcenia wynika, że

$$\varepsilon_1 E'_A = \varepsilon_2 E''_A. \qquad (19.19)$$

W rozważanym przypadku wektory natężenia pola były prostopadłe do powierzchni granicznej. W przypadku ogólnym, gdy wektor **E** ma kierunek ukośny względem

powierzchni granicznej, rozkładamy go na składową normalną E_n i styczną E_t. W stosunku do składowej normalnej obowiązuje równanie (19.19), a zatem

$$\varepsilon_1 E_{1n} = \varepsilon_2 E_{2n}, \tag{19.20}$$

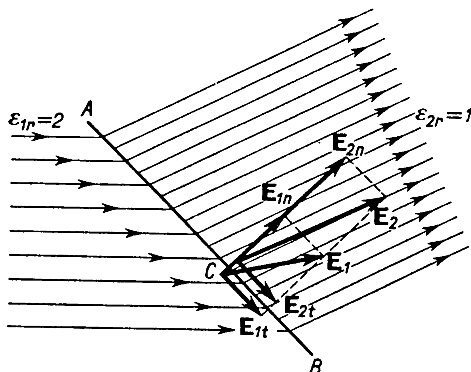
gdzie E_{1n} i E_{2n} przedstawiają składowe normalne natężenia pola na obu stronach powierzchni rozgraniczającej dwa dielektryki. Warto podkreślić, że (19.20) spełnione jest wtedy, gdy na powierzchni granicznej nie występują ładunki swobodne.

Z ostatniego równania wynika, że *składowe normalne natężenia pola przy przejściu przez powierzchnie graniczne rozdzielające dwa dielektryki zmieniają się odwrotnie proporcjonalnie do przenikalności elektrycznych*:

$$\frac{E_{1n}}{E_{2n}} = \frac{\varepsilon_2}{\varepsilon_1}.$$

Składowa styczna natężenia pola na granicy dwóch ośrodków nie ulega zmianie:

$$E_{1t} = E_{2t}. \tag{19.21}$$



Rys. 19.12

Otrzymane wnioski zastosujemy do przedstawienia graficznego linii sił i wektorów **E** na granicy dwóch dielektryków. Niech powierzchnia *AB* ukośna względem linii sił pola (rys. 19.12) oddziela ośrodek *I* o względnej przenikalności elektrycznej, np. $\varepsilon_{1r} = 2$ od ośrodka *II* o względnej przenikalności elektrycznej $\varepsilon_{2r} = 1$. W ośrodku pierwszym natężenie pola wynosi $\mathbf{E}_1$. Wektor $\mathbf{E}_1$ odkładamy w dowolnym punkcie *C* powierzchni granicznej. Rozkładamy go na dwie składowe $\mathbf{E}_{1t}$ i $\mathbf{E}_{1n}$ zgodnie z zasadą równoległoboku. W ośrodku drugim składowa $\mathbf{E}_{2n}$ będzie miała wartość dwa razy większą (odwrotna proporcjonalność do przenikalności elektrycznej). Składowa $\mathbf{E}_{2t}$ równa się $\mathbf{E}_{1t}$. Sumując geometrycznie wektory $\mathbf{E}_{2t}$ i $\mathbf{E}_{2n}$ znajdujemy wektor $\mathbf{E}_2$ różniący się zarówno wartością, jak i kierunkiem od wektora $\mathbf{E}_1$. Linie sił pola w ośrodku drugim kreślimy równolegle do $\mathbf{E}_2$. Gęstość ich rozmieszczenia jest większa od gęstości linii sił w ośrodku pierwszym. Krótko ujmujemy te wyniki stwierdzając, że na granicy dwóch dielektryków linie sił pola na ogół ulegają załamaniu i ich gęstość się zmienia.

19.6. Wektor indukcji elektrostatycznej

Wobec nieciągłości strumienia linii sił pola w ośrodkach niejednorodnych nie możemy w nich stosować prawa Gaussa w poznanej poprzednio postaci. Można jednak wprowadzić nowy wektor związany z **E**, ale tak dobrany, by nie ulegał zmianie na granicy dwóch ośrodków. Taką właściwość ma *wektor indukcji elektrostatycznej* **D** zdefiniowany równaniem

$$\mathbf{D} = \varepsilon \mathbf{E}^*. \tag{19.22}$$

Obliczmy przykładowo indukcję elektrostatyczną w odległości r od ładunku punktowego Q. Natężenie pola w takim punkcie wyraża się wzorem (19.13), a zatem

$$D = \frac{1}{4\pi}\frac{Q}{r^2}. \tag{19.23}$$

Jak widać, wartość indukcji elektrostatycznej nie zależy od właściwości ośrodka. Ze wzoru (19.23) wynika, że *jednostką indukcji elektrostatycznej jest* C/m². (Może uderzać fakt, do którego jeszcze wrócimy, że w tych samych jednostkach wyraża się gęstość powierzchniowa ładunków $\sigma = dQ/dS$.)

W odniesieniu do wektora indukcji wprowadzamy również pojęcia linii indukcji, strumienia indukcji ψ_D i gęstości strumienia indukcji. *Gęstość strumienia indukcji* jest to liczba linii indukcji przechodzących przez 1 m² powierzchni prostopadłej do wektora indukcji. Według umowy gęstość strumienia indukcji odpowiada liczbowo wartości **D** w danym punkcie.

Pozostałe definicje i wzory mają również postać analogiczną do omówionych w § 19.3. Tak np. *strumień indukcji przez elementarną powierzchnię dS* ustawioną ukośnie względem wektora **D** równa się

$$d\psi_D = D\,dS\cos\alpha = \mathbf{D}\cdot d\mathbf{S},$$

gdzie $d\mathbf{S}$ ma takie samo znaczenie jak we wzorze (19.15b), a kąt α jest kątem między normalną do $d\mathbf{S}$ i wektorem **D**.

Strumień indukcji przez dowolną powierzchnię S równa się wektorowej całce powierzchniowej rozciągniętej na powierzchnię S:

$$\psi_D = \int_S \mathbf{D}\cdot d\mathbf{S}.$$

Do strumienia indukcji przechodzącego przez zamkniętą powierzchnię można też stosować prawo Gaussa. Skoro wektor **D** jest ε razy większy od wektora **E**, to i całkowity strumień indukcji przez określoną powierzchnię jest ε razy większy od całkowitego strumienia elektrycznego przez tę powierzchnię. A zatem w odniesieniu do zamkniętej powierzchni, w której obszarze znajdują się ładunki $Q_1, Q_2, \ldots, Q_n$, obowiązuje prawo Gaussa w postaci

$$\psi_D = (Q_1+Q_2+Q_3+\ldots+Q_n) = \sum Q. \tag{19.24}$$

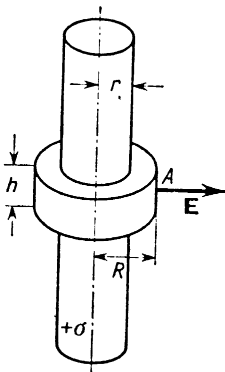
---

* Równanie to jest spełnione w ośrodkach izotropowych.

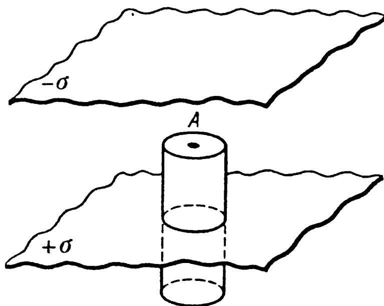
19.7. Zastosowanie prawa Gaussa

Znajomość prawa Gaussa ułatwia rozwiązywanie wielu zagadnień elektrostatycznych.

P r z y k ł a d 1. Obliczyć natężenie pola w punkcie A (rys. 19.13) znajdującym się w odległości R od osi nieskończenie długiego naładowanego walca o promieniu r i o gęstości powierzchniowej ładunku σ. Linie sił i linie indukcji skierowane są prostopadle do powierzchni bocznej walca.



Rys. 19.13



Rys. 19.14

Aby móc zastosować prawo Gaussa, wprowadźmy dodatkową powierzchnię zamkniętą, wykorzystującą możliwie maksymalnie symetrię pola elektrycznego w otoczeniu naładowanego walca. Niech to będzie np. powierzchnia drugiego walca o promieniu R i wysokości h, współosiowego z walcem pierwszym i usytuowanego na takiej wysokości, by punkt A znalazł się na jego powierzchni bocznej. Całkowity strumień indukcji ψ_D przez powierzchnię tego walca pomocniczego redukuje się do strumienia przez jego powierzchnię boczną (dlaczego?) o polu $S = 2\pi Rh$. Strumień ten można wyrazić jako

$$\psi_D = 2\pi RhD$$

lub — zgodnie z prawem Gaussa — jako

$$\psi_D = \sum Q = 2\pi rh\sigma.$$

Z porównania obu tych wyrażeń wynika, że

$$D = \frac{\sigma r}{R}.$$

A zatem szukane natężenie pola w punkcie A równa się

$$E = \frac{\sigma r}{\varepsilon R}. \qquad (19.25)$$

A więc w rozważanym przypadku natężenie pola jest wprost proporcjonalne do r i σ, a odwrotnie proporcjonalne do odległości R i przenikalności elektrycznej ośrodka, w którym znajduje się badany walec.

P r z y k ł a d 2. Obliczyć natężenie pola i indukcję elektrostatyczną w punkcie A przestrzeni między dwiema bardzo dużymi równoległymi płytami o stałej gęstości powierzchniowej ładunku odpowiednio równej $+\sigma$ i $-\sigma$. Punkt A leży daleko od krawędzi płyt, a jego odległość od płyt jest mała w porównaniu z ich rozmiarami. (Wyniki, jakie otrzymamy, można będzie stosować — przy zaniedbaniu efektów brzegowych — do kondensatorów płaskich o rozmiarach okładek dużych w stosunku do ich wzajemnej odległości.)

Z podobnych względów jak poprzednio wprowadzamy powierzchnię pomocniczą w postaci powierzchni walca przebijającego prostopadle jedną z płyt i tak ustawionego, by na jego górnej podstawie (rys. 19.14) przypadał rozważany punkt A. Wiemy już, że w przestrzeni między takimi płytami w dużych odległościach od ich krawędzi przebieg linii sił (i linii indukcji) jest taki jak na rys. 19.7. Linie są wzajemnie równoległe i prostopadłe do powierzchni płyt. A zatem strumień indukcji przez całą powierzchnię boczną walca, jak również przez jego podstawę dolną równa się zeru. Wchodzi zatem w grę jedynie strumień indukcji ψ_D przez górną podstawę walca, na której znajduje się punkt A. Strumień ten podobnie jak w poprzednim przykładzie można wyrazić dwojako:

$$\psi_D = SD$$

oraz

$$\psi_D = \sum Q = S\sigma,$$

gdzie S oznacza powierzchnię podstawy walca.

Z porównania obu wyrażeń wynika, że

$$D = \sigma, \tag{19.26a}$$

a stąd wynika dalej, że

$$E = \frac{\sigma}{\varepsilon}. \tag{19.26b}$$

Jak widać, wartość indukcji elektrostatycznej w dowolnym punkcie między dużymi płytami naładowanymi różnoimiennie, mającymi stałe gęstości powierzchniowe ładunku $+\sigma$ i $-\sigma$ jest równa tej gęstości powierzchniowej σ, a więc jednakowa we wszystkich punktach przestrzeni między płytami. Podobnie natężenie pola w całym obszarze między płytami jest jednakowe. Mamy zatem do czynienia z polem jednorodnym (zaburzenia tej jednorodności występują tylko przy brzegach płyt).

Warto tu raz jeszcze podkreślić bardzo istotny czynnik. W rozważanych przykładach tak wybieraliśmy pomocnicze zamknięte powierzchnie (tzw. powierzchnie Gaussa), by możliwie jak najbardziej wykorzystać symetrię pola elektrycznego, wytwarzanego przez obiekt badany. Taki warunek spełniała powierzchnia walcowa, obejmująca wycinek nieskończenie długiego, naładowanego pręta (przykład 1) i analogiczna powierzchnia ustawiona prostopadle do naładowanej płyty z przykładu 2. Brak symetrii wytwarzanego pola lub wybór nieodpowiedniej powierzchni pomocniczej mogą uniemożliwić wykorzystanie prawa Gaussa do wyznaczenia wartości E lub D, choć oczywiście prawo to stosuje się w każdym przypadku.

Z badań doświadczalnych, jak również z rozważań teoretycznych opartych na prawie Gaussa, można wyciągać pewne wnioski dotyczące naładowanych przewodników w warunkach elektrostatycznych. Przede wszystkim można stwierdzić, że w warunkach elektrostatycznych ładunek Q wprowadzony na przewodnik rozmieszcza się wyłącznie na jego zewnętrznej powierzchni. Rozkład ładunku — w warunkach osiągniętego stanu równowagi — jest taki, że wewnątrz przewodnika nie działają żadne siły elektryczne. Natężenie pola elektrycznego **E** wewnątrz przewodnika w warunkach elektrostatycznych równa się zeru. Rozkład ładunków na powierzchni izolowanego, naelektryzowanego przewodnika jest zależny od jego kształtu. Gdy rozkład ładunku na powierzchni przewodnika jest równomierny, to gęstość powierzchniowa ładunku σ jest stała. Taka sytuacja istnieje np. na powierzchni przewodników kulistych, jak również — z pominięciem warunków brzegowych — na powierzchni okładek kondensatora i na powierzchniach bocznych przewodników walcowych. Gęstość powierzchniowa ładunku jest większa na tych elementach zewnętrznej powierzchni przewodnika, które mają większą krzywiznę.

19.8. Praca przenoszenia ładunku w polu elektrycznym. Napięcie elektryczne

Ładunek $+Q_0$ wprowadzony do pola elektrycznego podlega działaniu sił elektrycznych. W przypadku ładunku swobodnego siły pola wykonują pracę przesuwając go wzdłuż linii siły. Praca dW wykonana na nieskończenie małym elemencie ds takiej drogi wynosi

$$dW = F ds.$$

Gdy przesunięcie nie odbywa się wzdłuż linii siły, lecz wzdłuż drogi tworzącej z kierunkiem siły kąt α, praca dW równa się

$$dW = F ds \cos\alpha = \mathbf{F} \cdot d\mathbf{s}.$$

Na drodze o skończonej długości s, np. od punktu A do punktu B, praca W_{AB} wyraża się całką liniową:

$$W_{AB} = \int_A^B \mathbf{F} \cdot d\mathbf{s}.$$

Uwzględniając (19.12) otrzymamy

$$W_{AB} = \int_A^B Q_0 \mathbf{E} \cdot d\mathbf{s} = Q_0 \int_A^B \mathbf{E} \cdot d\mathbf{s}. \tag{19.27}$$

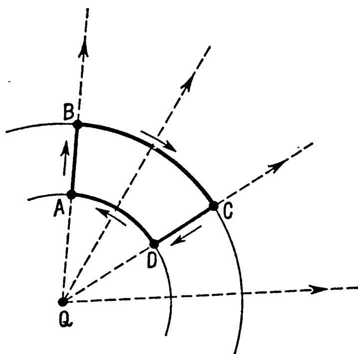
Wykonanie powyższej pracy nie jest związane z działaniem sił zewnętrznych, a zatem należy uważać, że praca ta jest wykonana kosztem zmniejszenia się energii potencjalnej, zmagazynowanej w polu elektrycznym. Zmiana wartości energii potencjalnej $E_{pA} - E_{pB}$ wynika ze zmiany konfiguracji ładunków w polu, a mianowicie z przesunięcia ładunku Q_0 z punktu A do punktu B. Możemy zatem napisać

$$W_{AB} = Q_0 \int_A^B \mathbf{E} \cdot d\mathbf{s} = E_{pA} - E_{pB}. \tag{19.28}$$

Gdybyśmy rozważali przesunięcie odwrotne (ruchem jednostajnym z prędkością nieskończenie małą) ładunku Q_0 z punktu B do punktu A, to wtedy praca $W_{BA}^{(z)}$ wykonana przez siły zewnętrzne (równoważące siły pola) zostałaby zużyta na wywołanie przyrostu energii potencjalnej $E_{pA}-E_{pB}$, czyli

$$W_{BA}^{(z)}=E_{pA}-E_{pB}. \tag{19.29}$$

Rozważając *pracę przenoszenia ładunku w polu elektrostatycznym po dowolnej krzywej zamkniętej* można wykazać, że *jest* ona *równa zeru.* Sprawdzimy to dla najprostszego przypadku przenoszenia ładunku Q_0 w polu ładunku punktowego Q po drodze $ABCDA$, zaznaczonej na rys. 19.15. Odcinki AB i CD tej drogi leżą na liniach sił pola, odcinki BC i DA — na łukach kół, które w każdym swym punkcie są prostopadłe do linii sił. Praca sił pola na odcinku AB jest równa co do wartości, lecz przeciwna co do znaku względem pracy wykonanej na odcinku CD. Prace na odcinkach BC i AD są równe zeru ze względu na prostopadłość kierunków siły i przesunięcia. A zatem całkowita praca na drodze zamkniętej $ABCDA$ jest równa zeru.



Rys. 19.15

Uogólnienie — już bez dowodu — otrzymanego wyniku na dowolne krzywe zamknięte pomyślane w polu elektrostatycznym pozwoli na wyciągnięcie wniosku o podstawowym znaczeniu: *siły pola elektrostatycznego są siłami zachowawczymi.* A zatem *praca tych sił nie zależy od kształtu drogi, lecz tylko od jej punktu początkowego i końcowego.* Analogiczne cechy przypisaliśmy siłom pola grawitacyjnego (por. § 4.6).

Stwierdzenie, że praca przenoszenia ładunków po drodze zamkniętej w polu elektrostatycznym równa się zeru, można matematycznie ująć w postaci

$$\oint \mathbf{E}\cdot d\mathbf{s}=0. \tag{19.30}$$

(Mamy tu do czynienia z całką liniową po dowolnej krzywej zamkniętej. Przypominamy, że symbol $d\mathbf{s}$ oznacza element drogi.)

Jest to równoznaczne ze stwierdzeniem, że w polu elektrostatycznym nie istnieją zamknięte linie sił. Pole takie nazywamy *polem bezwirowym.* Rzeczywiście, gdyby w polu elektrostatycznym istniały zamknięte linie sił, to ładunek przenoszony wzdłuż

całej takiej linii stale podlegałby działaniu siły zgodnej z kierunkiem ruchu, a więc wykonującej pracę dodatnią i całka $\oint \mathbf{E} \cdot d\mathbf{s}$ nie równałaby się zeru.

Poprzednio dowiedzieliśmy się, że praca przeniesienia ładunku Q_0 z punktu A do punktu B wyraża się wzorem (19.28). Obecnie już wiemy, że nie zależy ona od kształtu drogi łączącej punkt A z punktem B. Dzieląc wyrażenie (19.28) przez wartość przenoszonego ładunku Q_0 znajdziemy pracę przenoszenia ładunku jednostkowego z punktu A do punktu B. Będzie to wielkość charakterystyczna dla wybranych punktów pola A i B, zwana *napięciem elektrycznym* U_{AB} punktu A względem punktu B:

$$U_{AB} = \frac{W_{AB}}{Q_0} = \int_A^B \mathbf{E} \cdot d\mathbf{s} = \frac{E_{pA}}{Q_0} - \frac{E_{pB}}{Q_0}. \tag{19.31}$$

Jak widać z tego równania, jednostką napięcia elektrycznego w układzie SI jest J/C. Jednostkę tę nazywamy *woltem*.

$$1 \text{ V} = 1 \text{ J/C}.$$

Napięcie między dwoma punktami pola równa się jednemu woltowi, gdy praca przeniesienia ładunku jednego kulomba z jednego punktu do drugiego równa się jednemu dżulowi.

19.9. Potencjał elektryczny

Jak już wspominaliśmy w § 19.2, pole elektryczne można charakteryzować wektorowo za pośrednictwem natężenia pola $\mathbf{E}$ i skalarnie — za pośrednictwem wielkości zwanej potencjałem elektrycznym V.

Przez *potencjał elektryczny danego punktu pola* rozumiemy pracę, jaką muszą wykonać siły pola, aby ładunek jednego kulomba przenieść z danego punktu do punktu nieskończenie odległego (czasem mówimy — do nieskończoności) lub pracę, jaką muszą wykonać siły zewnętrzne przy przeniesieniu ładunku jednego kulomba z nieskończoności do danego punktu. Jako punkt odniesienia wybieramy punkt nieskończenie odległy, tj. taki, w którym siły pola równają się zeru. Z definicji potencjału wynika, że potencjał punktu nieskończenie odległego również równa się zeru.

Analizując podaną wyżej definicję potencjału stwierdzamy, że potencjał danego punktu A pola jest równocześnie napięciem tego punktu względem punktu nieskończenie odległego:

$$V_A = \frac{W_{A\infty}}{Q_0} = \frac{W^{(z)}_{\infty A}}{Q_0} = \frac{E_{pA}}{Q_0} - \frac{E_{p\infty}}{Q_0}.$$

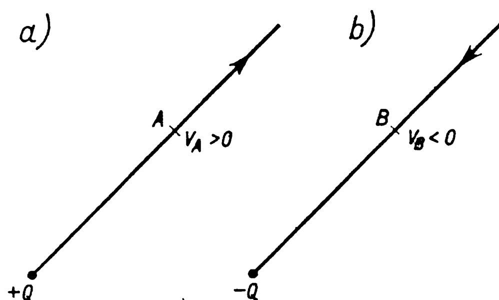
Wprowadzając umowę, że energia potencjalna ładunku w punkcie nieskończenie odległym równa się zeru otrzymujemy

$$V_A = \frac{E_{pA}}{Q_0}. \tag{19.32}$$

W oparciu o tę ostatnią zależność możemy *potencjał* w danym punkcie zdefiniować

jako stosunek energii potencjalnej ładunku Q_0 znajdującego się w tym punkcie do wartości Q_0 tego ładunku. Należy zatem dobrze pamiętać, że potencjał danego punktu pola i energia potencjalna ładunku w danym punkcie pola są to wielkości różne, związane ze sobą zależnością (19.32).

Jednostka potencjału jest oczywiście taka sama jak jednostka napięcia. *Wartość bezwzględna potencjału danego punktu równa się jednemu woltowi*, jeśli praca przeniesienia ładunku jednego kulomba z danego punktu do nieskończoności lub z nieskończoności do danego punktu równa się jednemu dżulowi.



Rys. 19.16

Potencjał danego punktu pola jest wielkością skalarną, która może przyjmować wartości zarówno dodatnie, jak i ujemne. Rozpatrzmy np. punkt A (rys. 19.16a) w otoczeniu ładunku dodatniego Q wytwarzającego pole. Przeniesienie ładunku dodatniego Q_0 równego jednemu kulombowi z punktu A do nieskończoności odbywa się pod działaniem sił pola kosztem malejącej do zera energii potencjalnej. A zatem w punkcie początkowym drogi, czyli w punkcie A energia potencjalna i potencjał są dodatnie. Ponieważ punkt A był wybrany dowolnie, więc można powiedzieć, że wszystkie punkty w otoczeniu odosobnionego ładunku punktowego dodatniego mają potencjały dodatnie.

I odwrotnie — przeniesienie ładunku dodatniego Q_0 równego jednemu kulombowi z punktu B (rys. 19.16b) w otoczeniu ładunku ujemnego $-Q$ wytwarzającego pole do nieskończoności wymaga wkładu pracy sił zewnętrznych na pokonanie przyciągania między różnoimiennymi ładunkami $-Q$ i $+Q_0$. Praca ta wywołuje przyrost energii potencjalnej przenoszonego ładunku od pewnej wartości początkowej do przyjętej umownie za zerową energii potencjalnej ładunku w nieskończoności. A zatem w punkcie początkowym drogi, czyli w punkcie B, istnieje ujemna energia potencjalna i ujemny potencjał. To samo dotyczy każdego punktu w otoczeniu odosobnionego ładunku punktowego ujemnego.

Warto może podkreślić, że niezależność pracy przenoszenia ładunku w polu elektrostatycznym od kształtu drogi decyduje o tym, że potencjał danego punktu pola jest jednoznacznie określony.

Łatwo sprawdzić, że

$$U_{AB} = V_A - V_B.$$

Wystarczy w tym celu wziąć np. pod uwagę drogę zamkniętą $AB\infty A$ i pracę na niej wykonaną przez siły pola, odniesioną do jednostki przenoszonego ładunku, przyrównać do zera:

$$\frac{W_{AB}}{Q_0}+\frac{W_{B\infty}}{Q_0}+\frac{W_{\infty A}}{Q_0}=0.$$

Pierwszy składnik tej sumy oznacza, zgodnie z definicją, napięcie U_{AB}, drugi — również zgodnie z definicją — oznacza V_B (gdyż bierzemy pod uwagę pracę sił pola), zaś trzeci —

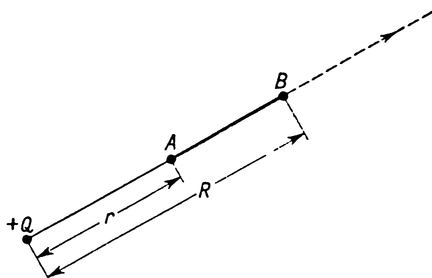
$$\frac{W_{\infty A}}{Q_0}=-\frac{W_{A\infty}}{Q_0}=-V_A.$$

A zatem

$$U_{AB}+V_B+(-V_A)=0,$$

$$U_{AB}=V_A-V_B. \tag{19.33}$$

19.9.1. *Potencjał w otoczeniu ładunku punktowego.* W celu wyprowadzenia wzoru na potencjał punktu pola w otoczeniu odosobnionego ładunku punktowego rozpatrzmy pracę przesunięcia ładunku Q_0 w polu ładunku dodatniego $+Q$ wzdłuż linii siły od punktu A (w odległości r od ładunku Q) do punktu B (w odległości R od ładunku Q) (rys. 19.17).



Rys. 19.17

Szukana praca wykonana przez siły pola jest równa

$$W_{AB}=\int_A^B F\,dr=\int_r^R \frac{QQ_0}{4\pi\varepsilon r^2}\,dr,$$

$$W_{AB}=\frac{QQ_0}{4\pi\varepsilon}\left(\frac{1}{r}-\frac{1}{R}\right). \tag{19.34}$$

Aby przejść do potencjału w punkcie A, otrzymane wyrażenie dzielimy przez Q_0 i zakładamy, że punkt B dąży do nieskończoności (tj. $R\to\infty$). Wtedy

$$V_A=\frac{W_{A\infty}}{Q_0}=\frac{Q}{4\pi\varepsilon}\frac{1}{r}=\frac{1}{4\pi\varepsilon}\frac{Q}{r}. \tag{19.35}$$

Potencjał w otoczeniu odosobnionego ładunku punktowego Q jest więc proporcjonalny do wartości tego ładunku, odwrotnie proporcjonalny do odległości badanego punktu od ładunku wytwarzającego pole i zależny od rodzaju ośrodka.

Energia potencjalna względem nieskończoności ładunku Q_0 umieszczonego w punkcie A równa się

$$E_{pA} = Q_0 V_A = \frac{Q_0 Q}{4\pi\varepsilon r}. \tag{19.36}$$

Uwzględniając (19.35) można (19.34) przepisać w postaci

$$W_{AB} = Q_0\left(\frac{Q}{4\pi\varepsilon r} - \frac{Q}{4\pi\varepsilon R}\right) = Q_0(V_A - V_B) = Q_0 U_{AB}. \tag{19.37}$$

A więc *praca przenoszenia ładunku Q_0 przez siły pola z punktu A do punktu B* równa się iloczynowi przenoszonego ładunku i różnicy potencjałów w obu punktach (lub iloczynowi przenoszonego ładunku i napięcia między punktami A i B).

Jeżeli pole wytwarzane jest przez wiele ładunków punktowych $Q_1, Q_2, Q_3, \ldots, Q_n$, to potencjał dowolnego punktu tego pola znajdujemy jako sumę algebraiczną potencjałów wytwarzanych przez poszczególne ładunki. Potencjał bowiem, jak już podkreślaliśmy, jest wielkością skalarną*.

19.9.2. *Powierzchnie ekwipotencjalne.* W otoczeniu ładunku punktowego potencjał zmienia się odwrotnie proporcjonalnie do odległości. Stąd wniosek, że wszystkie punkty leżące na kuli o promieniu r, w której środku znajduje się ładunek punktowy, mają ten sam potencjał. Wspomniana kula jest tzw. powierzchnią ekwipotencjalną. Ogólnie mówiąc, *powierzchnię ekwipotencjalną* tworzy zbiór punktów o jednakowej wartości potencjału. W odniesieniu do odosobnionych ładunków punktowych lub naładowanych przewodników kulistych powierzchnie ekwipotencjalne mają kształt kul. W polu układu ładunków punktowych, jak również w otoczeniu ciał naelektryzowanych niekulistych mogą one przyjmować najróżnorodniejszy kształt. Rys. 19.18 przedstawia układ powierzchni ekwipotencjalnych w otoczeniu dipola.

Ze wzoru (19.37) wynika, że praca przesuwania ładunku po powierzchni ekwipotencjalnej równa się zeru. Jest to możliwe dzięki temu, że linie sił są w każdym punkcie prostopadłe do powierzchni ekwipotencjalnych. Wynika to bezpośrednio z zależności

$$W = \mathbf{F}\cdot\mathbf{s} = Fs\cos\alpha = 0.$$

Ponieważ podczas przesuwania ładunku po powierzchni ekwipotencjalnej ani siła F, ani droga s nie równają się zeru, więc

$$\cos\alpha = 0, \quad \text{czyli} \quad \alpha = 90°.$$

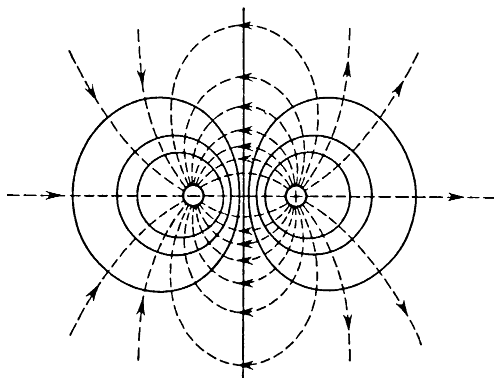
---

* Potencjał punktu pola elektrycznego wytworzonego przez naładowaną powierzchnię S wyraża się wzorem

$$V = \frac{1}{4\pi\varepsilon}\int_S \frac{\sigma dS}{r},$$

gdzie σ oznacza gęstość powierzchniową ładunku, r — odległość od elementarnego ładunku powierzchniowego σdS do badanego punktu pola. Całkowanie rozciągnięte jest na powierzchnię S.

Z tych rozważań wynika, że znajomość przebiegu linii sił (czyli znajomość natężenia pola w każdym punkcie pewnego obszaru) umożliwia wykreślenie prostopadłych do nich powierzchni ekwipotencjalnych i odwrotnie, znajomość przebiegu powierzchni ekwipotencjalnych (czyli znajomość potencjału w każdym punkcie pewnego obszaru) umożliwia narysowanie układu linii sił jako prostopadłych do tych powierzchni.



Rys. 19.18

Rozważając działanie sił na powierzchni naładowanego przewodnika w warunkach elektrostatycznych łatwo stwierdzić, że na takiej powierzchni możliwe jest tylko działanie sił elektrycznych, prostopadłych do powierzchni. Siły ukośne bowiem miałyby zawsze składowe styczne do powierzchni, a te wywoływałyby ruch ładunków, co byłoby sprzeczne z założeniem warunków elektrostatycznych. Stąd wniosek, że *powierzchnia przewodnika w warunkach elektrostatycznych jest powierzchnią ekwipotencjalną.*

P y t a n i e. Jak przedstawiają się powierzchnie ekwipotencjalne w jednorodnym polu? (Omówienie i rysunek.)

19.10. Zależność między natężeniem pola elektrostatycznego i potencjałem

Aby ustalić zależność między natężeniem pola elektrostatycznego $\mathbf{E}$ i potencjałem V, rozważmy pracę, jakiej wymaga przeniesienie ładunku Q_0 z punktu A leżącego na powierzchni ekwipotencjalnej o wartości potencjału V do punktu C na nieskończenie bliskiej powierzchni ekwipotencjalnej o wartości potencjału $V+dV$ (rys. 19.19). Ponieważ praca nie zależy od drogi, więc można założyć, że przenoszenie ładunku odbywa się po drodze ABC. Ale ruch ładunku wzdłuż powierzchni ekwipotencjalnej nie wymaga pracy, więc $W_{BC} = 0$ i cała praca przemieszczenia ładunku z A do C redukuje się do pracy przesunięcia go z A do B wzdłuż normalnej do obu nieskończenie bliskich powierzchni ekwipotencjalnych. Kierunek przesunięcia od A do B jest zgodny z kierunkiem, wzdłuż którego na jednostkę długości przypada największy wzrost potencjału $dV/ds = (dV/ds)_{\max}$, a przeciwny względem kierunku natężenia pola $\mathbf{E}$. Wobec założonej nieskończenie małej odległości ds można przyjmować, że wartość E na całej tej

drodze jest stała. Przesunięcie ładunku dodatniego na tej drodze ruchem jednostajnym wymaga zastosowania siły zewnętrznej $\mathbf{F} = -Q_0\mathbf{E}$. Praca tej siły

$$dW_{AB}^{(z)} = -Q_0 E ds.$$

Ale ta praca może być także wyrażona jako (por. (19.29) i (19.37)):

$$dW_{AB}^{(z)} = Q_0[(V+dV)-V],$$

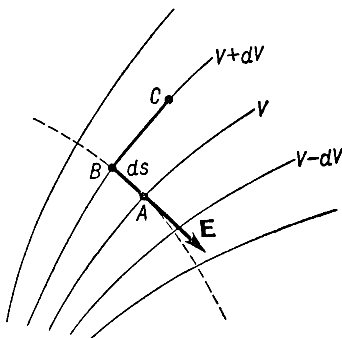
$$dW_{AB}^{(z)} = Q_0 dV.$$

Z porównania obu wzorów wynika, że

$$-Q_0 E ds = Q_0 dV,$$

skąd

$$E = -\frac{dV}{ds}. \tag{19.38}$$



Rys. 19.19

Ponieważ ds jest mierzone wzdłuż normalnej do powierzchni ekwipotencjalnych, czyli wzdłuż kierunku, dla którego $dV/ds = (dV/ds)_{max}$, więc można napisać

$$\mathbf{E} = -\operatorname{grad} V.$$

Innymi słowy, wartość liczbowa wektora **E** równa się różnicy potencjałów przypadającej na jednostkę długości, odmierzonej wzdłuż kierunku, dla którego $dV/ds = (dV/ds)_{max}$. Znak minus przypomina, że zwrot wektora **E** jest przeciwny do zwrotu najgwałtowniejszego wzrostu potencjału.

W polu elektrostatycznym jednorodnym na każdą jednostkę długości odmierzoną wzdłuż kierunku linii siły przypada stale ta sama różnica potencjałów. W takim polu obowiązuje zależność liczbowa

$$E = \frac{V_1 - V_2}{d} = \frac{U}{d}, \tag{19.39}$$

gdzie d oznacza najmniejszą odległość między punktami o potencjałach V_1 i V_2 (mierzoną wzdłuż linii siły).

Jak widać z ostatnich wzorów, *natężenie pola elektrycznego można wyrażać w* V/m

w układzie SI (por. § 19.2). Właśnie ta jednostka natężenia pola jest w najczęstszym użyciu.

Z tych zależności można wyciągnąć jeszcze jeden wniosek, a mianowicie: *potencjał w każdym punkcie przewodnika w warunkach elektrostatycznych jest taki sam zarówno wewnątrz przewodnika, jak i na jego powierzchni zewnętrznej*. Pamiętamy bowiem, że wewnątrz przewodnika natężenie pola elektrycznego **E** równa się zeru, a zatem i różnica potencjałów dV (por. wzór (19.38)) musi być równa zeru. Innymi słowy, dla każdej pary punktów przewodnika w warunkach elektrostatycznych spełniona jest równość

$$V_1 = V_2.$$

19.11. Pojemność elektryczna

Omówienie kolejnej wielkości elektrycznej, zwanej pojemnością, rozpoczniemy od kilku uwag ogólnych.

W zasadzie *pojęcie pojemności elektrycznej wiąże się z układem ciał*. Mówimy np. o pojemności układu złożonego z dwóch płytek metalowych (okładek) rozsuniętych na pewną odległość i rozdzielonych dowolnym dielektrykiem (kondensator płaski), o pojemności układu złożonego z dwóch koncentrycznych kul metalowych rozdzielonych dielektrykiem (kondensator kulisty) lub o pojemności współosiowych metalowych walców również rozdzielonych dielektrykiem (kondensator cylindryczny). Nawet w przypadku gdy mowa o pojemności odosobnionego przewodnika, należy uważać, że istnieje druga otaczająca go „okładka", oddzielona od badanego przewodnika dielektrykiem i mająca nieskończenie duże wymiary. Po naładowaniu odosobnionego przewodnika na tej odsuniętej do nieskończoności „okładce" pojawia się ładunek tzw. indukcyjny, równy co do wartości, lecz przeciwnego znaku, podobnie jak w rzeczywistym naładowanym kondensatorze. Potencjał tej odsuniętej do nieskończoności okładki — zgodnie z przyjętą umową — równa się zeru. W odniesieniu do przewodników odosobnionych czasem wprowadza się nazwę pojemności bezwzlędnej, co właśnie ma przypominać, że mowa jest o pojemności danego przewodnika względem drugiej okładki znajdującej się w nieskończoności.

Pojemność bezwzględna odosobnionego przewodnika definiowana jest jako stosunek ładunku zgromadzonego na przewodniku do jego potencjału:

$$C = \frac{Q}{V}. \tag{19.40}$$

Analogiczna definicja dotyczy pojemności układu przewodników tworzącego kondensator. Jeśli na okładkach zgromadzone są ładunki $+Q$ i $-Q$, a między okładkami istnieje różnica potencjałów, czyli napięcie równe U, to *pojemnością kondensatora* jest

$$C = \frac{Q}{U}. \tag{19.41}$$

Oba wzory są w gruncie rzeczy identyczne, bowiem w przypadku odosobnionego przewodnika jego potencjał V jest właściwie jego napięciem względem „okładki" znaj-

dującej się w nieskończoności, a więc mającej potencjał zerowy. I odwrotnie, jeśli w przypadku rzeczywistego kondensatora zastosujemy uziemienie jednej z okładek, tj. potencjał tej okładki sprowadzimy do zera, to U we wzorze (19.41) będzie równe liczbowo potencjałowi okładki nieuziemionej.

Z równań (19.40) i (19.41) wynika definicja jednostki pojemności. *Jednostką pojemności w układzie* SI *jest farad* (F). *Farad* jest to pojemność takiego kondensatora, który naładowany ładunkiem jednego kulomba wykazuje różnicę potencjałów jednego wolta między okładkami:

$$1\,\mathrm{F} = 1\,\mathrm{C/V}.$$

Jest to jednostka bardzo duża. Często w użyciu są jej podwielokrotne, jak mikrofarad (10^{-6} F) i pikofarad (10^{-12} F).

Wzory (19.40) i (19.41) można też interpretować w ten sposób, że potencjał odosobnionego przewodnika (napięcie między okładkami kondensatora) jest wprost proporcjonalny do ładunku zgromadzonego na tym przewodniku (na okładce kondensatora). Współczynnikiem proporcjonalności jest pojemność bezwzględna odosobnionego przewodnika lub pojemność kondensatora. Jak zobaczymy niżej, jest to wielkość charakterystyczna dla danego układu. O wartości liczbowej pojemności danego układu decydują wymiary geometryczne okładek, ich odległość wzajemna oraz przenikalność elektryczna dielektryka przedzielającego okładki.

Zanim przejdziemy do bardziej szczegółowego omawiania kondensatorów, obliczymy pojemność bezwzględną odosobnionego przewodnika kulistego.

W oparciu o prawo Gaussa można wykazać (por. zad. 36), że przewodząca, odosobniona, naładowana ładunkiem Q kula o promieniu R wywołuje w swoim otoczeniu ($r \geqslant R$) takie pole elektryczne $\mathbf{E}$ i taki potencjał V, jak ładunek punktowy Q umieszczony w jej środku. A zatem potencjał na powierzchni kuli ($r = R$) wyraża się wzorem

$$V = \frac{1}{4\pi\varepsilon}\frac{Q}{R}.$$

Stąd pojemność bezwzględna odosobnionej kuli równa się

$$C = \frac{Q}{V} = 4\pi\varepsilon R.$$

W próżni, a w przybliżeniu i w powietrzu, pojemność takiej kuli jest równa

$$C = 4\pi\varepsilon_0 R.$$

Z tych wzorów widać wyraźnie zależność pojemności bezwzględnej odosobnionego przewodnika od jego rozmiarów (R) i od przenikalności elektrycznej ośrodka, w jakim on się znajduje.

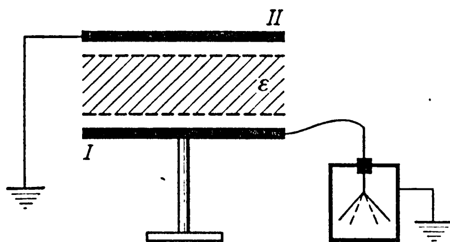
Pytanie. Ile wynosi pojemność elektryczna kuli ziemskiej traktowanej jako przewodnik odosobniony? Średni promień Ziemi wynosi 6400 km.

19.12. Kondensatory

Najwięcej uwagi poświęcimy kondensatorom płaskim, o których już zresztą kilkakrotnie mówiliśmy (por. § 19.3 i 19.7).

W celu wyprowadzenia wzoru na pojemność kondensatora płaskiego weźmy pod uwagę układ (rys. 19.20) złożony z dwóch płyt metalowych o powierzchniach S, rozsuniętych na odległość d, przedzielonych dielektrykiem o przenikalności elektrycznej ε. Jedna z płyt jest uziemiona, ma więc potencjał zerowy. Na drugiej płycie znajduje się ładunek Q; jej potencjał wynosi V. Między płytami istnieje zatem napięcie U, równe liczbowo potencjałowi V, związane z natężeniem E jednorodnego pola między okładkami zależnością (por. wzór (19.39))

$$E = \frac{U}{d}.$$



Rys. 19.20

Z zastosowania prawa Gaussa do kondensatora płaskiego (por. § 19.7) otrzymaliśmy

$$E = \frac{\sigma}{\varepsilon},$$

gdzie σ oznacza gęstość powierzchniową ładunku na każdej z okładek równą Q/S.

Z porównania obu tych wzorów wynika, że

$$\frac{U}{d} = \frac{Q}{\varepsilon S}.$$

Stąd pojemność kondensatora płaskiego

$$C = \frac{Q}{U} = \frac{\varepsilon S}{d} = \frac{\varepsilon_0 \varepsilon_r S}{d}. \tag{19.42}$$

Jest więc ona zależna od powierzchni okładek S, od ich wzajemnej odległości d oraz od przenikalności elektrycznej ε dielektryka przedzielającego okładki.

Kondensator próżniowy ma pojemność

$$C_0 = \frac{\varepsilon_0 S}{d}.$$

A zatem *względna przenikalność elektryczna* ε_r dielektryka przedzielającego okładki równa się

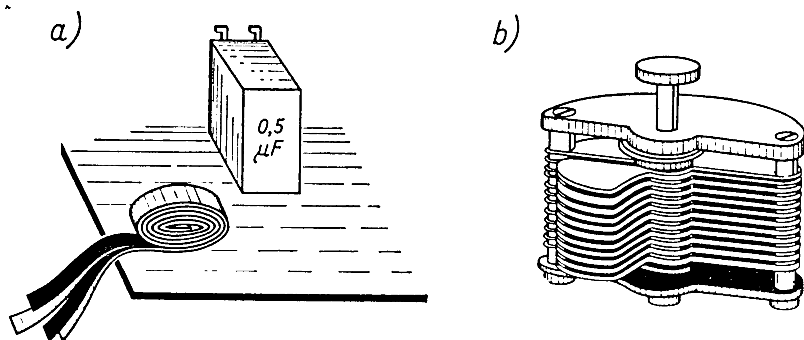
$$\varepsilon_r = \frac{C}{C_0}.$$

Jest to więc stosunek pojemności kondensatora, gdy jego okładki są przedzielone

danym dielektrykiem, do pojemności tegoż kondensatora, gdy między jego okładkami znajduje się próżnia.

Zależność pojemności kondensatora płaskiego od odległości okładek i od rodzaju dielektryka przedzielającego okładki można wykazać w prostym doświadczeniu. Jedną z okładek kondensatora ładujemy i łączymy z listkami elektroskopu*, którego osłona jest uziemiona (rys. 19.20). Drugą okładkę uziemiamy.

Rozchylenie listków elektroskopu jest miarą potencjału pierwszej okładki względem ziemi, a więc także miarą jej napięcia *U* względem okładki drugiej.



Rys. 19.21

Odsuwanie płyty drugiej wywołuje wzrost rozchylenia listków, czyli wzrost napięcia *U*, co — przy niezmiennym ładunku na okładkach — świadczy o zmniejszaniu się pojemności.

Wsuwanie między okładki kondensatora (pierwotnie powietrznego), umieszczone w stałej odległości *d*, dielektryków o coraz większej przenikalności elektrycznej ε wywołuje zmniejszanie się rozchylenia listków elektroskopu, czyli zmniejszanie się napięcia *U*. Świadczy to o wzroście pojemności.

Na zakończenie dorzucimy kilka uwag dotyczących kondensatorów często stosowanych w technice.

Techniczne kondensatory niskonapięciowe robi się często z pasków folii cynowej lub glinowej, przedzielonych papierem parafinowym (rys. 19.21a). Paski te można zwinąć tak, że zajmują mało miejsca, a jednak zapewniają dużą pojemność dzięki dużej powierzchni. Uzyskaniu dość znacznej pojemności sprzyja też stosunkowo duża wartość przenikalności elektrycznej papieru parafinowego.

Znane też są ogólnie kondensatory powietrzne o zmiennej pojemności. Ich budowę przedstawia rys. 19.21b.

Ostatnio stosuje się dość często jako dielektryki przedzielające okładki kondensatora masy ceramiczne z dodatkiem TiO_2 o przenikalności elektrycznej około 100.

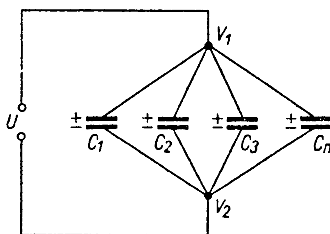
---

* Budowa elektroskopu laboratoryjnego znana jest z kursu szkolnego. Pamiętać należy, że rozchylenie listków elektroskopu mierzy w zasadzie różnicę potencjałów między listkami i osłoną. W przypadku osłony uziemionej rozchylenie listków mierzy ich potencjał względem Ziemi.

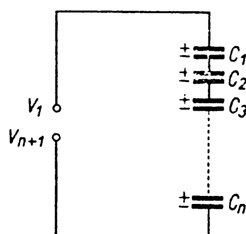
Kondensatorami o niewielkich wymiarach a dużej pojemności są kondensatory elektrolityczne. Można je otrzymać przepuszczając prąd elektryczny przez obwód złożony z dwóch elektrod glinowych zanurzonych w roztworze wodnym boraksu i kwasu borowego H_2BO_4. W czasie elektrolizy powstaje na anodzie glinowej warstwa nieprzewodzącego $Al(OH)_3$. Wobec małej grubości warstwy tego dielektryka pojemność otrzymanego w ten sposób kondensatora jest bardzo duża.

Kondensatory znalazły bardzo szerokie zastosowanie, toteż niejednokrotnie będziemy jeszcze o nich wspominali. Tutaj wymienimy ich zastosowanie w elektronice w rozmaitego typu obwodach generatorów, detektorów, wzmacniaczy, w optyce elektronowej (mikroskop elektronowy) oraz w technice wysokich napięć jako źródła silnych impulsów energetycznych (znaczne energie potencjalne są nagromadzone w polu między okładkami) itp.

Wprowadzenie dielektryka między okładki kondensatora nie tylko zwiększa pojemność, lecz także zwiększa — w stosunku do kondensatora powietrznego — granicę naładowania kondensatora, wywołującą tzw. elektryczne *przebicie*. W każdym dielektryku, gdy wartości natężenia pola elektrycznego **E** przekraczają pewne wartości graniczne, daje się zaobserwować przepływ prądu po przypadkowych liniach łamanych zależnych od struktury dielektryka, przyjmujący — przy dalszym wzroście **E** — postać iskrowego przebicia elektrycznego. Graniczna wartość natężenia pola **E**, powyżej której występują wspomniane zjawiska, nosi nazwę *wytrzymałości dielektryka na przebicie*. Dla powietrza w pobliżu ziemi przyjmuje się orientacyjną wartość wytrzymałości na przebicie 30 kV/cm. Względne wytrzymałości w stosunku do powietrza wynoszą np. dla ebonitu 2–3, gumy 2,5–5, oleju transformatorowego 1,2–7, papieru 1,5–3, porcelany 8–12, szkła 3–6.



Rys. 19.22



Rys. 19.23

Pytanie. Na dwóch płaskich kondensatorach o jednakowych wymiarach: 1) podtrzymujemy tę samą różnicę potencjałów między okładkami, 2) dostarczamy im tych samych ładunków. Jeden z kondensatorów jest kondensatorem próżniowym, a drugi jest wypełniony dielektrykiem o względnej przenikalności elektrycznej ε_r. Jaki jest w przypadku pierwszym stosunek ładunków na obu kondensatorach, w przypadku drugim — stosunek różnic potencjałów na okładkach?

Łączenie kondensatorów. Rozróżniamy łączenie kondensatorów 1) równoległe (rys. 19.22) i 2) szeregowe (rys. 19.23).

1. *Łączenie równoległe kondensatorów* polega na połączeniu przewodnikiem okładek jak na rys. 19.22. W ten sposób powstaje nowy układ o powiększonej powierzchni okładek. Wszystkie okładki połączone z biegunem dodatnim źródła mają potencjał V_1,

a wszystkie połączone z biegunem ujemnym — potencjał V_2. A zatem napięcie między okładkami każdego z kondensatorów składowych równa się U. Ładunki Q przypadające na poszczególne kondensatory można więc wyrazić jako

$$Q_1 = C_1 U$$
$$Q_2 = C_2 U$$
$$Q_3 = C_3 U$$
$$\dots\dots\dots$$
$$Q_n = C_n U.$$

Sumując stronami znajdziemy

$$Q_1+Q_2+Q_3+ \dots +Q_n = (C_1+C_2+C_3+ \dots +C_n)U.$$

Z porównania z ogólnym wzorem, wiążącym ładunek z napięciem między okładkami:

$$Q = CU,$$

wynika, że *układ równolegle połączonych kondensatorów ma pojemność równą sumie pojemności kondensatorów składowych*:

$$C = C_1+C_2+C_3+ \dots +C_n = \sum_{i=1}^{n} C_i. \qquad (19.43)$$

2. *Łączenie szeregowe kondensatorów* przedstawia rys. 19.23. Na okładkę pierwszą połączoną z biegunem dodatnim źródła o potencjale V_1 wprowadzamy ładunek $+Q$, ostatnią zaś łączymy z biegunem źródła o potencjale V_{n+1}. Na pozostałych (pośrednich) okładkach tworzą się przez indukcję ładunki odpowiednio $-Q$ i $+Q$. Napięcia między poszczególnymi okładkami będą równe:

dla pierwszej pary $V_1 - V_2 = \frac{Q}{C_1}$,

dla drugiej pary $V_2 - V_3 = \frac{Q}{C_2}$,

. . . .

dla n-tej pary $V_n - V_{n+1} = \frac{Q}{C_n}$.

Sumując stronami znajdziemy

$$V_1 - V_{n+1} = Q\left(\frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \dots + \frac{1}{C_n}\right),$$

gdzie $V_1 - V_{n+1} = U$ jest napięciem przyłożonym między okładką pierwszą i ostatnią.

Porównując ostatnie wyrażenie z zasadniczym wzorem wiążącym Q, U i C wypisanym w postaci

$$U = Q\frac{1}{C}$$

stwierdzamy, że *odwrotność pojemności układu złożonego z kondensatorów połączonych szeregowo równa się sumie odwrotności pojemności poszczególnych kondensatorów*:

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \ldots + \frac{1}{C_n}. \tag{19.44}$$

Łatwo sprawdzić, że taki układ kondensatorów połączonych szeregowo ma pojemność mniejszą od najmniejszej pojemności kondensatora wchodzącego w jego skład. Połączenie szeregowe ma zastosowanie przede wszystkim wtedy, gdy chcemy zmniejszyć możliwość uszkodzenia kondensatora przez przebicie.

19.13. Elektryczna energia potencjalna

Rozważania związane z elektryczną energią potencjalną rozpoczniemy od najprostszego przypadku odosobnionego przewodnika o pojemności bezwzględnej C. Naładowanie tego przewodnika do potencjału równego V związane jest z dostarczeniem mu ładunku Q_1 i wymaga wykonania pewnej pracy. Praca włożona jest miarą energii potencjalnej, wynikającej z istnienia na przewodniku ładunku Q_1 i towarzyszącego mu pola elektrycznego. A zatem

$$W = E_p = \int_0^{Q_1} V dQ = \int_0^{Q_1} \frac{Q}{C} dQ = \frac{1}{2}\frac{Q_1^2}{C}.$$

Korzystając z zależności między wielkościami Q, C i V można przedstawić elektryczną energię potencjalną naładowanego przewodnika w trzech równoważnych postaciach:

$$E_p = \frac{1}{2}\frac{Q_1^2}{C} = \frac{1}{2}CV^2 = \frac{1}{2}Q_1 V. \tag{19.45}$$

Przejdźmy do innego przypadku. Niech w pewnym punkcie przestrzeni istnieje ładunek punktowy Q_1 wytwarzający dokoła siebie pole, a w nieskończoności — drugi ładunek punktowy Q_2. Przenieśmy ładunek Q_2 z nieskończoności do punktu B odległego o r od ładunku Q_1, a więc o potencjale

$$V_B = \frac{1}{4\pi\varepsilon}\frac{Q_1}{r}.$$

Utworzenie takiego układu ładunków Q_1 i Q_2 wymaga pracy $W_{\infty B}$ związanej z przeniesieniem ładunku Q_2 z nieskończoności do punktu B:

$$W_{\infty B} = Q_2 \frac{1}{4\pi\varepsilon}\frac{Q_1}{r} = \frac{1}{4\pi\varepsilon}\frac{Q_1 Q_2}{r}.$$

Łatwo sprawdzić, że wynik byłby taki sam, gdybyśmy ładunek Q_1 przenieśli z nieskończoności do punktu odległego o r od ładunku Q_2. Pracę $W_{\infty B}$ lub $W_{\infty A}$ nazywamy *elektryczną energią potencjalną* E_p układu dwóch ładunków umieszczonych w odległości wzajemnej r:

$$E_p = \frac{1}{4\pi\varepsilon}\frac{Q_1 Q_2}{r}. \tag{19.46}$$

Zmiana odległości wzajemnej ładunków Q_1 i Q_2 wywołuje zmniejszenie lub wzrost elektrycznej energii potencjalnej układu.

P y t a n i e. Jak można wyznaczyć elektryczną energię potencjalną układu trzech ładunków punktowych umieszczonych w skończonych odległościach od siebie?

Podobnie obliczamy *elektryczną energię potencjalną kondensatora*. Obliczenia E_p sprowadzamy tym razem do obliczenia pracy całkowitego rozładowania kondensatora, czyli pracy przeniesienia całkowitego ładunku Q_0 z jednej okładki na drugą:

$$E_p = E_{\text{kond}} = W_{\text{rozład}} = \int_0^{Q_0} V\,dQ = \int_0^{Q_0} \frac{Q}{C}\,dQ,$$

$$E_{\text{kond}} = \tfrac{1}{2}\frac{Q_0^2}{C} = \tfrac{1}{2}CU^2. \tag{19.47}$$

Dla kondensatora płaskiego — uwzględniając wzory (19.39) i (19.42) — otrzymamy

$$E_{\text{kond płas}} = \tfrac{1}{2}\frac{\varepsilon S}{d}E^2 d^2.$$

Ale iloczyn Sd przedstawia objętość V dielektryka między okładkami kondensatora, a zatem

$$E_{\text{kond płas}} = \frac{\varepsilon E^2}{2}V. \tag{19.48}$$

Innymi słowy, elektryczna energia potencjalna kondensatora jest proporcjonalna do przenikalności elektrycznej ośrodka przedzielającego okładki, do kwadratu natężenia pola oraz do objętości dielektryka między okładkami. Można więc uważać, że właśnie ten dielektryk jest siedliskiem energii. Obliczamy ilość energii przypadającą na jednostkę objętości dielektryka między okładkami kondensatora płaskiego, czyli tzw. *gęstość elektrycznej energii potencjalnej w*:

$$w = \frac{E_{\text{kond płas}}}{V} = \frac{\varepsilon E^2}{2}. \tag{19.49}$$

Można wykazać, że wyprowadzony wzór na gęstość elektrycznej energii potencjalnej jest wzorem ogólnym: gęstość elektrycznej energii potencjalnej dowolnego pola elektrostatycznego w otoczeniu punktu, w którym natężenie pola równa się **E**, wyraża się wzorem

$$w = \frac{dE_{\text{pola}}}{dV} = \frac{\varepsilon E^2}{2}.$$

Znając wzór na elektryczną energię potencjalną kondensatora płaskiego można wyznaczyć siłę wzajemnego przyciągania jego okładek. Załóżmy, że odległość między okładkami zmniejszyła się o ds. Towarzyszy temu zmiana energii $dE_{\text{kond płas}}$ równa

$$dE_{\text{kond płas}} = \frac{\varepsilon E^2}{2}S\,ds.$$

Ale ta zmiana energii równa jest pracy mechanicznej siły F na drodze ds, a więc

$$\frac{\varepsilon E^2}{2}S\,ds = F\,ds,$$

skąd

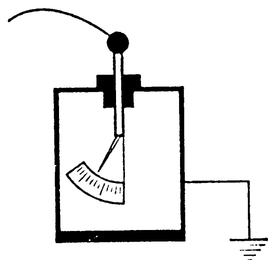
$$F = \frac{\varepsilon E^2}{2} S = wS. \tag{19.50}$$

Innymi słowy, *siła wzajemnego przyciągania okładek kondensatora płaskiego jest proporcjonalna do kwadratu natężenia pola między okładkami, do przenikalności elektrycznej dielektryka i do powierzchni okładek* albo — w innym ujęciu — jest proporcjonalna do gęstości elektrycznej energii potencjalnej kondensatora płaskiego i do jego powierzchni.

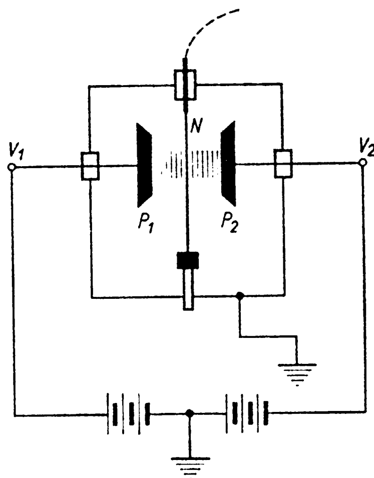
P y t a n i e. Czy energia kondensatora płaskiego powietrznego naładowanego do różnicy potencjałów V_0 zwiększy się czy zmniejszy (i ile razy) przy wprowadzeniu między okładki dielektryka o przenikalności elektrycznej ε?

19.14. Elektrometry

Nazwę elektrometrów nadajemy licznym przyrządom, które po odpowiednim wycechowaniu służą do pomiaru ładunku lub potencjału elektrycznego. Najprostszymi elektrometrami są wspominane już poprzednio wyskalowane elektroskopy jedno- lub dwulistkowe (rys. 19.24).



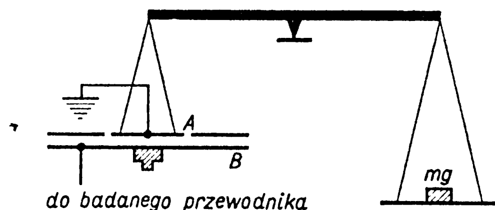
Rys. 19.24



Rys. 19.25

Dokładniejszym elektrometrem jest elektrometr strunowy (rys. 19.25). Nić platynowa N o średnicy 1–2 μ zwisa słabo naciągnięta między dwiema płytami P_1 i P_2 naelektryzowanymi do potencjałów V_1 i V_2 równych co do wartości, lecz o przeciwnych znakach. Całość znajduje się w uziemionej osłonie. Nić połączona jest z przewodnikiem o badanym potencjale. W zależności od znaku i wielkości tego potencjału nić wygina się bardziej lub mniej w stronę płytki P_1 lub P_2. Po wycechowaniu przyrządu z wielkości wychylenia wnioskujemy o badanym potencjale.

Innym typem elektrometru jest elektrometr bezwzględny. Zasadę jego budowy przedstawia rys. 19.26. Widzimy tam czułą wagę, w której zamiast jednej z szalek zasto-



Rys. 19.26

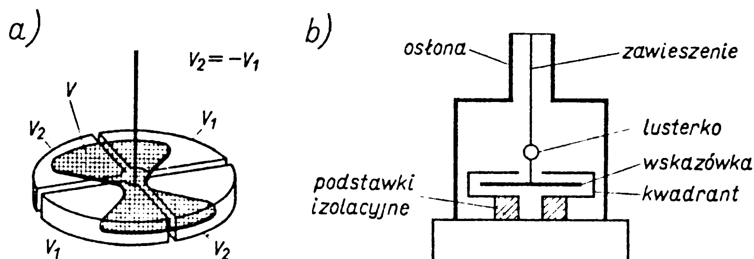
sowana jest uziemiona płyta kondensatora płaskiego A. Wagę doprowadzamy do równowagi, a następnie drugą okładkę kondensatora, czyli płytę B łączymy z badanym przewodnikiem. Obie płyty znajdują się w odległości d od siebie i działają na siebie siłą F wyrażoną równaniem (19.50). Chcąc zapobiec zbliżeniu się szalki A do B, musimy siłę F zrównoważyć odpowiednim ciężarem mg, czyli

$$mg = \frac{\varepsilon E^2}{2} S$$

lub — uwzględniając (19.39) —

$$mg = \frac{\varepsilon V^2}{2d^2} S.$$

Pomiar szukanego potencjału elektrycznego sprowadzony jest w tej metodzie do wyznaczania ciężaru mg (w niutonach) dodatkowego odważnika, określenia powierzchni płyty kondensatora S (w m^2) i wzajemnej odległości między płytami d (w m).



Rys. 19.27

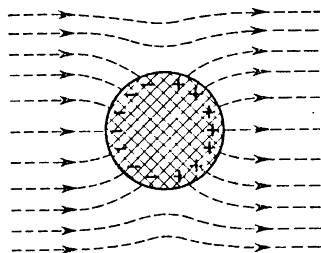
Na innej zasadzie opiera się działanie elektrometru kwadrantowego (rys. 19.27). Podstawową częścią tego elektrometru jest umieszczony w osłonie na podstawkach izolacyjnych wydrążony cylinder, o wysokości ok. 1 cm i średnicy ok. 5 cm, podzielony na cztery części, zwane kwadrantami. Wewnątrz cylindra, równolegle do jego podstaw, wisi lekka wskazówka, np. z glinu, w kształcie ósemki. Może ona być zawieszona na cienkiej kwarcowej nici o grubości kilku mikronów, pokrytej cienką warstwą srebra. Kwadranty połączone są ze sobą parami na krzyż i mają równe co do wartości, lecz o przeciwnych znakach potencjały ($V_2 = -V_1$). Wskazówka połączona jest z przewodnikiem, którego potencjał badamy. W takich warunkach wskazówka poddana jest

działaniu pary sił wywołującej obrót. Nić kwarcowa skręca się o pewien kąt, który, jak wykazuje teoria, jest proporcjonalny do mierzonego potencjału. Można więc po wyskalowaniu przyrządu bezpośrednio z wielkości kąta skręcenia wnioskować o wartości potencjału. Pomiar kąta skręcenia odbywa się metodą optyczną, gdyż do nici przymocowane jest małe zwierciadełko płaskie (opis metody por. § 24.8).

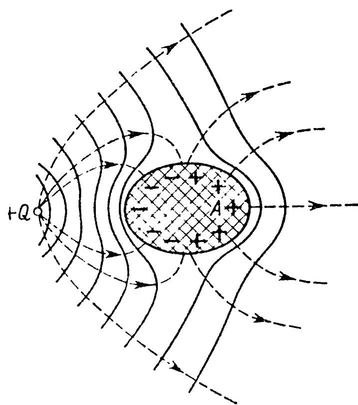
Przy pomiarach potencjału za pomocą elektrometru kwadrantowego stosuje się także tzw. idiostatyczny układ połączeń, polegający na tym, że wskazówka i jedna para kwadrantów są uziemione, a druga para jest połączona z badanym przewodnikiem. Przy tym sposobie łączenia — jak wykazuje teoria — kąt obrotu wskazówki, a więc i zwierciadełka, jest w przybliżeniu proporcjonalny do kwadratu mierzonego potencjału.

19.15. Przewodniki w polu elektrostatycznym. Elektryzowanie przez indukcję

O natężeniu pola elektrycznego i potencjale elektrycznym przewodników już mówiliśmy. Obecnie zajmiemy się jeszcze krótko elektryzowaniem przewodników przez indukcję. Przypomnimy sytuację panującą na metalicznych okładkach kondensatora płaskiego. Okładka *A* (rys. 19.7) naelektryzowana np. dodatnio wywołuje uporządkowanie



Rys. 19.28



Rys. 19.29

ładunków na okładce *B*. Ładunki ujemne w postaci swobodnych elektronów grupują się w nadmiarze na powierzchni okładki *B* zwróconej do *A*. Na przeciwległej powierzchni okładki *B* (lub na ziemi — w przypadku uziemienia okładki *B*) gromadzi się nadmiar ładunków dodatnich. Linie sił pola wychodzą z ładunków dodatnich na okładce *A* i kończą się na ładunkach ujemnych na okładce *B*. Z teorii kondensatora płaskiego (por. wzór (19.26a)) wiadomo, że gęstość powierzchniowa ładunków indukowanych na okładce związana jest z indukcją elektrostatyczną *D*:

$$D = \sigma.$$

Gdyby ładunki indukowane powstawały na powierzchni ukośnej względem linii sił pola elektrycznego, to obowiązywałaby zależność

$$D_n = \sigma,$$

gdzie D_n oznacza wartość składowej wektora D w kierunku normalnej do powierzchni.

Umieszczenie przewodnika nienaładowanego w polu wywołuje zaburzenie przebiegu linii sił pola. Na rys. 19.28 przedstawione jest zaburzenie jednorodności pola, wynikające z powstawania ładunków indukowanych na powierzchni kulistego przewodnika. Linie sił kończą się i zaczynają na powierzchni przewodnika. Wewnątrz przewodnika nie ma pola. Rysunek 19.29 przedstawia zmianę przebiegu (pierwotnie radialnych) linii sił pola (linie kreskowane) i zmianę kształtu (pierwotnie kulistych) powierzchni ekwipotencjalnych (linie ciągłe) w otoczeniu ładunku punktowego dodatniego, wywołane wprowadzeniem do pola przewodnika A.

19.16. Dielektryki w polu elektrostatycznym

Zanim przejdziemy do omówienia zachowania się dielektryków w polu elektrostatycznym, zatrzymamy się chwilę nad sprawami ogólniejszymi.

Atom, jak wiemy, składa się z dodatnio naładowanego jądra i z ujemnie naładowanych elektronów otaczających jądro. Całkowity ładunek dodatni jądra jest równy liczbowo sumie ładunków elektronów. Dzięki swemu szybkiemu ruchowi elektrony działają na otoczenie tak, jakby ich ładunek przypadał w środku ich orbit, przyjmowanych w pierwszym przybliżeniu za kołowe. W ten sposób to „średnie" położenie elektronów, które nazywać będziemy centrum ładunków ujemnych, pokrywa się z centrum ładunków dodatnich jądra. W oddziaływaniu zewnętrznym nie ujawnia się ani ładunek atomu, ani moment dipolowy. Atom znajdujący się poza obrębem pola elektrycznego jest tworem obojętnym pod względem elektrycznym.

Gdy jednak atom znajdzie się w polu elektrycznym, siły pola mogą spowodować przesunięcie centrum ładunków dodatnich w jedną stronę, a centrum ładunków ujemnych — w przeciwną. W wyniku tych przesunięć otrzymujemy dipol o ładunkach $+Q$ i $-Q$, rozsuniętych na odległość a. Dipol taki ma moment elektryczny p,

$$p = Qa,$$

skierowany, według umowy, od ładunku ujemnego do ładunku dodatniego wzdłuż osi dipola. Efekt powyższy nazywamy *polaryzacją elektronową*.

Przejdźmy teraz do rozważania tworów bardziej skomplikowanych od atomów — zajmijmy się cząsteczkami.

Okazuje się, że dielektryki mogą się składać z cząsteczek niepolarnych lub cząsteczek polarnych (dipolowych). Do *cząsteczek niepolarnych* zaliczamy cząsteczki o budowie symetrycznej, np. cząsteczki dwuatomowych gazów H_2, O_2 itp., cząsteczki typu CH_4, CCl_4 itp. W tych cząsteczkach — podobnie jak w atomach — centrum ładunków dodatnich pokrywa się z centrum ładunków ujemnych. Bez działania zewnętrznego pola cząsteczki tego typu nie wykazują właściwości elektrycznych: ich moment dipolowy $\mathbf{p} = 0$. Pod wpływem pola centra obu rodzajów ładunków rozsuwają się — mó-

wimy wówczas o *polaryzacji elektronowej cząsteczki* lub o *polaryzacji deformacji*. Rozsunięciu centrów ładunków towarzyszy powstanie indukowanego dipola o momencie p proporcjonalnym do E. Po usunięciu pola moment dipolowy wraca do wartości zerowej.

Do *cząsteczek polarnych* zaliczamy cząsteczki wody, alkoholi, eterów, a przede wszystkim cząsteczki utworzone z jonów dodatnich i ujemnych. Cząsteczki polarne mają już „z natury" rozsunięte centra ładunków dodatnich i ujemnych, a zatem i bez działania zewnętrznego pola każda z cząsteczek ma moment dipolowy $\mathbf{p} \neq 0$ (por. tab. 19.1). Jednak w dużym zbiorowisku cząsteczek polarnych nie poddanych działaniu pola układ tych molekularnych dipoli jest najzupełniej chaotyczny i wypadkowy moment dipolowy równa się $\sum \mathbf{p} = 0$. Po umieszczeniu dielektryka polarnego w polu elektrostatycznym występuje tendencja do uporządkowania ustawienia dipoli — do przyjęcia przez momenty dipolowe kierunku zgodnego z kierunkiem pola. To skręcanie się poszczególnych cząsteczek w polu wywołuje efekt zwany *polaryzacją dipolową dielektryka lub polaryzacją orientacji*. Towarzyszyć jej może także dodatkowe rozsunięcie centrów ładunków dodatnich i ujemnych, czyli polaryzacja elektronowa.

Tabela 19.1

Momenty dipolowe

Rodzaj cząsteczki	p 10^{-30}C · m
woda	6,1
alkohole, etery	~6
kwasy organiczne	~4,7
ketony	~9
rodnik C—O	~0,23
rodnik C=O	~7,7

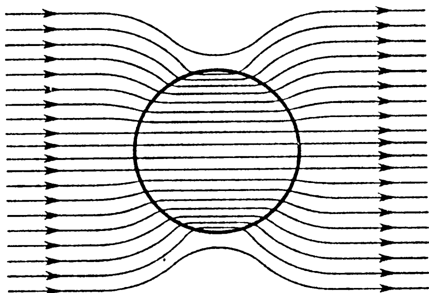
Chaotyczne ruchy cieplne cząsteczek są czynnikiem utrudniającym uzyskanie pełnego uporządkowania orientacji dipoli. One też sprawiają, że po usunięciu pola dielektryk wraca do stanu bezładnego rozmieszczenia cząsteczek, do stanu niespolaryzowania. Możemy jednak polaryzację dielektryka utrwalić przez „zamrożenie". Na przykład dielektryk w postaci mieszaniny żywicy i wosku w stanie roztopionym umieszczony w polu elektrycznym ulega polaryzacji dipolowej. Jeżeli przed usunięciem pola obniżymy temperaturę tak, że mieszanina się zestali, ruchy cieplne będą utrudnione i stan polaryzacji dielektryka utrzyma się trwale. Taki elektryczny odpowiednik trwałego magnesu nazwano *elektresem*.

Przesunięcia ładunków w dielektrykach odbywają się tylko w obrębie atomu lub cząsteczki. Dlatego też w dielektryku nie można rozdzielić ładunków dodatnich i ujemnych np. przez odprowadzenie ładunków jednego znaku do ziemi. W tym tkwi różnica elektryzowania indukcyjnego przewodników i polaryzacji dielektryków.

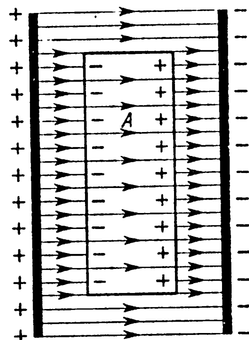
W paragrafie 19.15 mówiliśmy już o deformacji pola wywołanej umieszczeniem przewodnika w polu elektrycznym (rys. 19.28 i 19.29). Obecnie zajmiemy się deformacją pola wywołaną przez dielektryk.

Na rysunku 19.30 przedstawiony jest przebieg linii indukcji pierwotnie jednorodnego pola elektrycznego po wprowadzeniu do niego kuli z dielektryka o przenikalności większej od otoczenia. Widać zachowanie ciągłości linii indukcji, ale też wyraźnie zaznacza się zmiana gęstości ich rozmieszczenia. Wzrost gęstości linii w dielektryku zachodzi kosztem zmniejszenia gęstości linii w najbliższym otoczeniu.

Charakter deformacji zależy od właściwości, kształtu, rozmiarów i rozkładu dielektryków w polu, jak również od rozkładu ładunków wywołujących pole. Ponieważ siły oddziaływania między ładunkami maleją ze wzrostem przenikalności elektrycznej ośrodka przedzielającego, więc i natężenie pola maleje w tym samym stosunku.



Rys. 19.30



Rys. 19.31

Rozpatrzmy bardziej szczegółowo przykład następujący: w jednorodnym polu płaskiego kondensatora próżniowego umieszczono płytkę A z dielektryka (rys. 19.31). Nastąpiło wtedy uporządkowanie dipoli naturalnych lub wytworzenie dipoli w polu elektrycznym lub też i jedno i drugie. Wynik jest taki, że na obu powierzchniach płytki A zwróconych do okładek kondensatora pojawiły się ładunki różnoimienne $+Q$ i $-Q$. Nazywać je będziemy indukowanymi ładunkami polaryzacji dielektryka. Mimo tej polaryzacji płytka jako całość jest elektrycznie obojętna. Na rys. 19.31 przedstawiony jest przebieg linii sił pola w dielektryku i jego otoczeniu. Tym razem widzimy nieciągłość przebiegu linii sił na granicy dielektryk-próżnia, gdyż rysunek odnosi się do linii natężenia pola elektrycznego, a nie do linii indukcji.

Należy sobie dobrze uświadomić, że w zasadzie mamy do czynienia ze *zjawiskiem polaryzacji* (*uporządkowanie molekularnych dipoli*) *zachodzącym w całej objętości dielektryka*, czyli z efektem objętościowym. Uzewnętrznia się on jednak za pośrednictwem powierzchniowo rozłożonych *ładunków indukowanych* polaryzacji.

Rozważając działanie dielektryka na otoczenie możemy pójść jeszcze o krok dalej. Zamiast zajmować się efektem objętościowym jego polaryzacji, można brać pod uwagę równoważny mu efekt, wynikający z działania ładunków zastępczych w takiej ilości i tak rozmieszczonych w przestrzeni, jak były rozłożone na powierzchni dielektryka

indukowane ładunki polaryzacji. Innymi słowy, można sobie wyobrazić, że z pola został usunięty dielektryk, a zamiast niego w próżni pojawiły się ładunki $+Q$ i $-Q$. Rozłożone są one w przestrzeni dokładnie w ten sposób, jak rozmieszczone były indukowane ładunki polaryzacji na powierzchniach płytki A zwróconych do okładek kondensatora i prostopadłych do linii sił pola podstawowego. Równoważność polaryzacji dielektryka i tych ładunków sprowadza się do tego, że moment dipolowy przypadający na jednostkę objętości dielektryka (czyli tzw. *wektor polaryzacji* **P**), wynikający z efektu polaryzacji dielektryka, musi być równy momentowi dipolowemu wywołanemu przez ładunki zastępcze i przypadającemu na jednostkę objętości tej przestrzeni, która poprzednio była zajęta przez dielektryk. Moment dipolowy wywołany przez ładunki zastępcze równa się Qd, gdzie d oznacza grubość płytki A. Objętość obszaru między tymi ładunkami równa się Sd, gdzie S oznacza powierzchnię płytki A. A zatem

$$P = \frac{Qd}{Sd} = \frac{Q}{S} = \sigma.$$

Znaleźliśmy w ten sposób związek między wartością liczbową wektora polaryzacji **P** i gęstością powierzchniową σ indukowanych ładunków polaryzacji:

$$P = \sigma. \qquad (19.51)$$

Ta prosta zależność obowiązuje wtedy, gdy związane ładunki polaryzacji są zgrupowane równomiernie na powierzchni dielektryka, prostopadłej do kierunku natężenia pola elektrycznego **E**. Wtedy wektor **P** jest równoległy do wektora **E**.

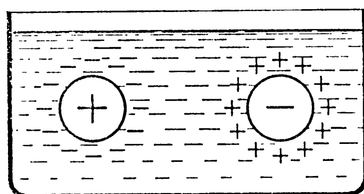
Wpływ dielektryka na pole, do którego jest on wniesiony, można też ująć inaczej. Wiemy już, że możemy w myśli usunąć płytkę dielektryka z pola, pozostawiając tylko w próżni w niezmienionych położeniach wytworzone poprzednio na płytce indukowane ładunki polaryzacji, spełniające warunek (19.51). Ładunki te są źródłem nowych linii sił, które w dawnym obszarze płytki A (stanowiącym obecnie obszar próżniowy) biegną w kierunku przeciwnym do kierunku linii podstawowego pola $\mathbf{E}_0$ kondensatora. Są to więc linie pola $\mathbf{E}'$ odwrotnie skierowanego niż pole $\mathbf{E}_0$. A zatem wypadkowe natężenie pola w obszarze zajętym przez dielektryk jest różnicą natężeń pola podstawowego $\mathbf{E}_0$ (między okładkami kondensatora próżniowego) i pola $\mathbf{E}'$, wynikającego z polaryzacji dielektryka.

Nasze rozważania dotyczyły kondensatora płaskiego: stwierdziliśmy osłabienie pola wynikające z wprowadzenia dielektryka, co oczywiście pociąga za sobą zmniejszenie różnicy potencjałów między okładkami. Wynik ma jednak znaczenie ogólne: w dielektryku wprowadzonym do pola elektrycznego powstaje pole elektryczne o natężeniu mniejszym od natężenia pola pierwotnego.

Zajmiemy się obecnie wynikami doświadczalnych badań dielektryków.

1. Wiemy już, że przeniesienie ładunków Q_1 i Q_2 z próżni do jakiegoś dielektryka powoduje zmniejszenie siły wzajemnego oddziaływania w stosunku odpowiadającym wartości względnej przenikalności elektrycznej danego ośrodka, przy zachowaniu odległości między ładunkami. Jakościowe wytłumaczenie tego faktu jest łatwe: dzięki polaryzacji dielektryka ładunek dodatni zostaje otoczony ładunkami ujemnymi na odpo-

wiednio ustawionych dipolach molekularnych, a dokoła ładunku ujemnego grupują się ładunki dodatnie dipoli molekularnych (rys. 19.32). Obecność tych ładunków przeciwnych znaków w najbliższym otoczeniu ładunków Q_1 i Q_2 powoduje zmniejszenie siły ich wzajemnego oddziaływania.



Rys. 19.32

2. Wiele badań poświęcono przenikalności elektrycznej. Przegląd podstawowych wiadomości o przenikalności elektrycznej zaczniemy od podania tabeli wartości liczbowych ε_r różnych dielektryków (tab. 19.2).

Tabela 19.2

Względne przenikalności elektryczne ε_r

Gazy warunki normalne		Ciała stałe temperatura pokojowa		Ciecze temperatura pokojowa	
Hel	1,000074	papier	~2–3	benzen	2,3
H_2	1,00026	ebonit	2–3	aceton	21,5
O_2	1,00055	parafina	~2	alkohol metylowy	33,7
N_2	1,00061	drewno	2–8	nitrobenzen	36,1
powietrze	1,00059	siarka	4		
CO_2	1,00098	szkło sodowe	4–6	woda	81,1
para wodna	1,0078	szkło ołowiowe	10	alkohol etylowy	25,7
para alkoh. etyl.	1,0065				
		porcelana	6,5		
		polistyren	2,6		
		alkohol etylowy zestalony	3,1		

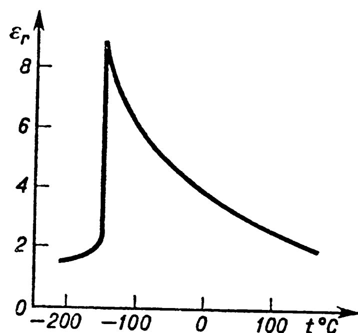
W pierwszej części tabeli podane są wartości względnej przenikalności elektrycznej kilku gazów. Jak widać, wartości ε_r są bliskie jedności. Dane tabeli (z wyjątkiem dotyczących pary wodnej i pary alkoholu etylowego) odnoszą się do gazów w warunkach normalnych. Ze wzrostem ciśnienia przenikalność elektryczna gazów rośnie, ze wzrostem temperatury — maleje.

W przypadku ciał stałych różnice względnych przenikalności elektrycznych poszczególnych ciał są znacznie większe niż w przypadku ciał gazowych. Tutaj — w przeciwieństwie do ciał gazowych — dużą rolę odgrywa natura elektryczna cząsteczek, a mianowicie ich naturalna polarność i indukowana polaryzowalność.

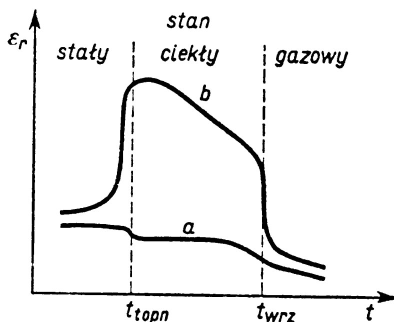
Interesujące jest zestawienie wartości względnych przenikalności elektrycznych dla danego ciała w stanie stałym i ciekłym:

lód	2,85	woda	81,7
alkohol etylowy zestalony	3,1	alkohol etylowy ciekły	25,7

Przejściu ze stanu ciekłego do gazowego i ze stanu stałego do ciekłego towarzyszy skok wartości przenikalności elektrycznej. Widać to z wykresu (rys. 19.33) odnoszącego się do eteru etylowego. „Skok" ε_r odpowiada temperaturze topnienia. Rząd wielkości tych skoków zależny jest od tego, czy cząsteczka jest polarna, czy nie. Ogólne pojęcie o charakterze tych zmian daje rys. 19.34. Krzywa *a* wykresu odnosi się do cząsteczek niepolarnych, krzywa *b* — do cząsteczek polarnych. Wzrost przenikalności elektrycznej przy topnieniu ciał, mających cząsteczki polarne, tłumaczy się większą swobodą ruchów molekularnych dipoli. W substancjach niepolarnych brak jest trwałych dipoli skręcających się do położenia równoległego do kierunku pola.



Rys. 19.33



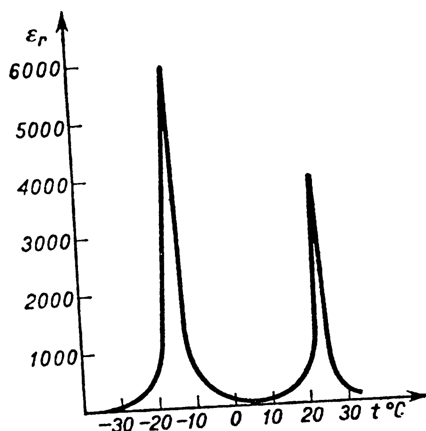
Rys. 19.34

Zwiększanie temperatury bez przejścia przez zmianę fazy prowadzi do zmniejszenia polaryzacji ciekłych dielektryków zarówno polarnych, jak i niepolarnych, z tym tylko że w pierwszym przypadku zmiany przenikalności elektrycznej są dość duże, w drugim — nieznaczne.

Na uwagę zasługuje odrębna grupa ciał, tzw. ferroelektryków, wyróżniających się bardzo dużymi wartościami względnych przenikalności elektrycznych w pewnych przedziałach temperatur. Do grupy tej należą: sól Seignette'a (winian sodowo-potasowy o wzorze $NaKC_4H_6O_6 \cdot 4H_2O$), kwaśny fosforan potasu KH_2PO_4, tytanian baru $BaTiO_3$ itp.

Wykres na rys. 19.35 wskazuje, jak przebiegają zmiany przenikalności elektrycznej soli Seignette'a w przedziale temperatur $-30 - +30$°C w słabym polu elektrycznym. Widać z wykresu, że w tym przedziale temperatur względna przenikalność elektryczna dwukrotnie osiąga bardzo duże wartości rzędu tysięcy. Kwaśny fosforan wapnia ma dużą względną przenikalność elektryczną w przedziale temperatur między -190°C i -130°C. Tytanian baru ma w temperaturze 20°C względną przenikalność elektryczną równą 1000–2000, a przy wprowadzeniu pewnych domieszek ε_r dochodzi do 8000.

Ferroelektryki mają pewne właściwości analogiczne do właściwości ferromagnetyków, a mianowicie ich przenikalność elektryczna zależy od natężenia pola elektrycznego, po usunięciu pola wykazują pozostałość elektryczną, a przy zmianach natężenia pola — zjawisko histerezy.



Rys. 19.35

Na zakończenie jeszcze jedna uwaga. W przypadku niektórych krystalicznych dielektryków polaryzację można wywołać działaniami mechanicznymi lub termicznymi. Tak np. ściskanie lub rozciąganie odpowiednio wyciętych płytek kwarcowych powoduje występowanie ładunków dodatnich i ujemnych na ich powierzchniach. Efekt ten nazywamy *efektem piezoelektrycznym*. Omówiliśmy go nieco szczegółowiej przy omawianiu ultradźwięków (por. § 12.7). Okazuje się, że gwałtowne ogrzewanie lub oziębianie krystalicznych dielektryków również powoduje polaryzację. To zjawisko nazywamy *efektem piroelektrycznym*.

19.17. Wyznaczanie ładunku elementarnego metodą Millikana

Najmniejszy ładunek elektryczny — „atom" elektryczności, zwany *ładunkiem elementarnym*, jest to ładunek pojedynczego elektronu, a tym samym ładunek jednowartościowego jonu. Jest to podstawowa wielkość fizyczna wyznaczana różnymi metodami.

Przeanalizujmy dokładniej zasadę pomiaru ładunku elementarnego w metodzie Millikana (rys. 19.36). W komorze K utrzymujemy stałą temperaturę (termostatowanie) i niezmienne ciśnienie, którego stałość sprawdzamy na manometrze M. W komorze umieszczamy kondensator płaski C z otworem w górnej okładce. Nad tym otworem znajduje się wylot rozpylacza, z którego wydmuchujemy drobne kropelki oliwy o promieniu rzędu 1 μm. Wskutek tarcia o ścianki rozpylacza krople ulegają naelektryzowaniu. Krople opadają i część z nich dostaje się między okładki kondensatora. Pojedynczą kroplę odpowiednio oświetloną obserwujemy w obszarze między okładkami

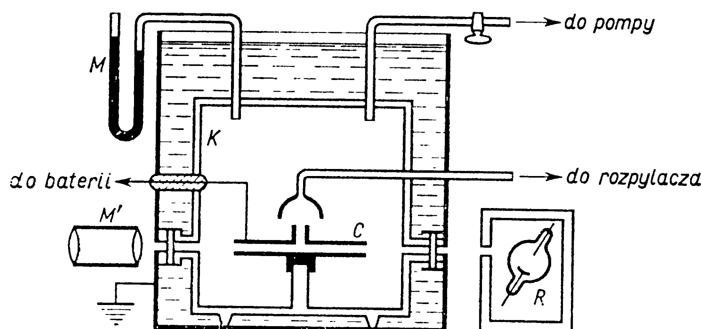
przez mikroskop M'. Zależnie od znaku ładunku kropli Q oraz kierunku i wartości natężenia E pola elektrycznego może być ona przyspieszana lub opóźniana, przesuwana w górę lub w dół. Możliwe jest też utrzymywanie kropli w spoczynku w dowolnym położeniu. Ten ostatni przypadek wymaga zrównoważenia ciężaru kropli mg przez siłę elektrostatyczną QE i siłę wyporu F_w

$$mg = QE+F_w$$

lub

$$mg = Q\frac{U}{d}+\tfrac{4}{3}\pi r^3 \varrho_{pow} g, \tag{19.52}$$

gdzie U oznacza napięcie między okładkami, d — ich odległość, r — promień kropli, ϱ_{pow} — gęstość powietrza w danych warunkach. Do wyznaczenia ładunku Q z tego



Rys. 19.36

równania (przy znanych U, d, ϱ_{pow} i g) brakuje nam wartości masy m i promienia r kropli. Znając gęstość oliwy ϱ_{ol} w danych warunkach obliczylibyśmy masę, gdybyśmy znali promień kropli. Jest on jednak zbyt mały, by można było zmierzyć go bezpośrednio. Metoda pośrednia polega na obserwacji zachowania się kropli oliwy między okładkami kondensatora przy braku pola elektrycznego. W tym przypadku kropla opada ku dołowi pod wpływem własnego ciężaru mg. Po uzyskaniu pewnej prędkości v ruch kropli staje się jednostajny dzięki zrównoważeniu się siły ciężkości z siłą oporu lepkości (wyrażoną prawem Stokesa) i siłą wyporu (wyrażoną prawem Archimedesa), czyli

$$mg = 6\pi\eta r v+\tfrac{4}{3}\pi r^3 \varrho_{pow} g$$

lub

$$\tfrac{4}{3}\pi r^3 \varrho_{ol} g = 6\pi\eta r v+\tfrac{4}{3}\pi r^3 \varrho_{pow} g, \tag{19.53}$$

gdzie η oznacza współczynnik lepkości powietrza w danych warunkach. Po obliczeniu r z równania (19.53) znajdujemy masę kropli i z równania (19.52) obliczamy Q.

Pomiary przeprowadzane wielokrotnie przez Millikana wykazały, że ładunki kropelek oliwy, które można było zmieniać stosując naświetlanie promieniami Röntgena (por. § 29.8) z lampy R są zawsze niewielkimi wielokrotnościami ładunku $e = 1{,}60\times\times 10^{-19}$ C. Jest to właśnie wartość ładunku elementarnego.

19.18. Podsumowanie wiadomości

Na zakończenie tego dość trudnego, a równocześnie bardzo podstawowego rozdziału podamy przykładowe podsumowanie przerobionego w nim materiału*.

W elektrostatyce kolejno zajmowaliśmy się:

1) ładunkiem elektrycznym,

2) wytwarzanym przez niego polem elektrycznym,

3) opisem pola elektrycznego:

a) wektorowym — za pomocą wektora natężenia pola **E** i wektora indukcji elektrostatycznej **D** (prawo Gaussa dostarcza powiązania **E** lub **D** za pośrednictwem strumienia ψ_E lub ψ_D z ładunkiem Q, wytwarzającym pole),

b) skalarnym — za pomocą potencjału elektrycznego V (związek $\mathbf{E} = -\operatorname{grad} V$),

4) pojemnością elektrostatyczną,

5) kondensatorami,

6) wyznaczaniem ładunku elementarnego.

Poznane wielkości zestawiamy w tabelce, uwzględniając równanie definicyjne wielkości, jej jednostkę w układzie SI oraz stosowane oznaczenie literowe.

Wielkość	Oznaczenie wielkości	Równanie definicyjne	Jednostka w układzie SI	Oznaczenie jednostki
ładunek elektryczny	Q	$Q = \int_{t_1}^{t_2} I\,dt$	kulomb	C
natężenie pola elektrycznego	**E**	$\mathbf{E} = \frac{\mathbf{F}}{Q_0}$	niuton na kulomb	N/C
		$\mathbf{E} = -\operatorname{grad} V$	wolt na metr	V/m
indukcja elektryczna (przesunięcie dielektryczne)	**D**	$\mathbf{D} = \varepsilon\mathbf{E}$	kulomb na metr kwadratowy	C/m^2
potencjał	V	$V_A = \frac{W_{A\infty}}{Q_0}$	wolt	V
napięcie	U	$U = V_A - V_B$	wolt	V
pojemność elektrostatyczna	C	$C = \frac{Q}{V}$	farad	F

Do podstawowych poznanych praw należą: prawo Coulomba i prawo Gaussa (podać ich dokładne sformułowania).

Do podstawowych wzorów należą wzory wyrażające: pole elektrostatyczne w otoczeniu ładunku punktowego, pracę przenoszenia ładunku w polu elektrycznym, potencjał w otoczeniu ładunku punktowego, pojemność kondensatora płaskiego, pojemność układów kondensatorów, energię potencjalną elektryczną. Byłoby wskazane ułożenie planu wyprowadzenia ważniejszych wzorów.

* Proponujemy samodzielne opracowanie analogicznych podsumowań po innych rozdziałach.

Pytania i zadania

1. Czym się zajmuje elektrostatyka?
2. Czy ładunki elektryczne występują w przyrodzie tylko tam, gdzie istnieją ciała naelektryzowane?
3. Co można przyjmować za kryterium podziału ciał na przewodniki, półprzewodniki i izolatory?
4. Jakie prawo określa siłę wzajemnego oddziaływania dwóch ładunków elektrycznych z odległości r? Jakie cechy ma powyższa siła?
5. Jaką wartość liczbową ma stała dielektryczna próżni ε_0 i co wyraża względna stała dielektryczna dowolnego ośrodka?
6. Kiedy mówimy, że w określonym obszarze przestrzeni istnieje pole elektrostatyczne? Jaka jest definicja natężenia pola elektrostatycznego i jego jednostki w układzie SI? Co rozumiemy przez linie sił pola? Podać przykłady przebiegu linii sił pola w różnych przypadkach. Czy z gęstości przebiegu linii sił można wnioskować o natężeniu pola?
7. Czy natężenie pola jest wielkością skalarną, czy wektorową?
8. Jak obliczamy natężenie pola pochodzącego od dwóch ładunków punktowych?
9. Co rozumiemy przez strumień elektryczny i w jakich jednostkach go mierzymy?
10. Jaka jest treść prawa Gaussa? Podać przykłady jego zastosowania. Wyjaśnić dlaczego w przypadku rurek sił: a) strumień elektryczny przez powierzchnię boczną równa się zeru, b) całkowity strumień przez powierzchnię rurki równa się zeru?
11. Dlaczego napotykamy trudności przy stosowaniu prawa Gaussa, jeśli zamknięta powierzchnia pomocnicza obejmuje granicę dwóch dielektryków?
12. Jakie zależności obowiązują na granicy dwóch dielektryków w odniesieniu do składowych a) stycznych i b) normalnych natężenia pola elektrycznego. Jaki jest przebieg linii sił pola na granicy dwóch dielektryków?
13. Jak usunięto trudności w stosowaniu prawa Gaussa do ośrodków niejednorodnych?
14. Jaka jest definicja wektora indukcji **D** i jego strumienia?
15. Jaką wielkość wykorzystuje się do skalarnego opisu pola elektrycznego? Podać dokładną definicję tej wielkości i jej jednostki w układzie SI.
16. Wyjaśnić dlaczego w definicji potencjału elektrycznego jest mowa o punkcie leżącym w nieskończoności? Czy wybór tego punktu odniesienia jest w zasadzie dowolny?
17. Jak wyprowadza się wzór na potencjał dowolnego punktu pola w otoczeniu ładunku punktowego?
18. Co rozumiemy przez powierzchnię ekwipotencjalną? Jak są względem siebie rozmieszczone linie sił i powierzchnie ekwipotencjalne?
19. Czemu się równa w warunkach elektrostatycznych: a) natężenie pola i potencjał wewnątrz naładowanego przewodnika, b) potencjał we wszystkich punktach jego powierzchni.
20. Jak można wyprowadzić podstawową zależność między natężeniem pola i spadkiem potencjału przypadającym na jednostkę długości?
21. Jaka zależność istnieje między ładunkiem wprowadzonym na dany przewodnik i jego potencjałem? Jaki współczynnik łączy we wzorze obie wspomniane wielkości? Podać definicję tego współczynnika i jego jednostki w układzie SI.
22. Jakie urządzenia noszą nazwę kondensatorów? Jak można uzasadnić tę nazwę? Jakie korzyści wynikają ze stosowania kondensatorów?
23. Omówić budowę różnych rodzajów kondensatorów i sposoby ich łączenia. Jak zmienia się pojemność układu kondensatorów w zależności od sposobu ich połączenia?
24. Omówić elektryczną energię potencjalną w przypadku a) naładowanego izolowanego przewodnika, b) układu złożonego z dwóch ładunków punktowych, c) naładowanego kondensatora.
25. Jaka jest zależność między gęstością elektrycznej energii potencjalnej i natężeniem pola?
26. Omówić zasady działania podstawowych elektrometrów.
27. Jakie są skutki wprowadzenia do istniejącego pola elektrycznego a) nienaładowanego przewodnika, b) dielektryka?
28. Omówić budowę niepolarnych i polarnych cząsteczek. Czy wpływ zewnętrznego pola na oba

rodzaje cząsteczek jest taki sam? Zdefiniować moment dipolowy cząsteczki. Scharakteryzować polaryzację elektronową i polaryzację orientacji.

29. Jak się zmieniają przenikalności elektryczne dielektryków polarnych i niepolarnych w zależności od temperatury?

30. Co to jest ładunek elementarny? Jaka jest jego wartość liczbowa? Na czym polega metoda Millikana jego wyznaczania?

31. Dwa ładunki dodatnie Q i nQ znajdują się w odległości a. W jakiej odległości od mniejszego ładunku należy umieścić trzeci ładunek, aby działająca na niego siła wypadkowa była równa zeru?

Odp. $a(\sqrt{n}-1)/(n-1)$.

32. Dwie jednakowo naładowane kulki o jednakowych masach m zawieszono na jedwabnych niciach o jednakowej długości L przymocowanych do jednego punktu. Obliczyć ładunki na kulkach, jeśli wiadomo, że kulki odepchnęły się na odległość r. (Założenie upraszczające: $L \gg r$.)

Odp. $\sqrt{4\pi\varepsilon_0 mgr^3/2L}$.

33. Na jaką odległość rozsunęłyby się dwie kulki mające każda ładunek Q, promień R i gęstość ϱ_1, zawieszone na niciach jedwabnych jednakowej długości L zamocowanych w jednym punkcie, po zanurzeniu ich w nafcie o gęstości ϱ_2 i względnej przenikalności elektrycznej ε_r. (Założenie upraszczające: L znacznie większe od szukanej odległości.)

34. Jakie jest natężenie pola i potencjał w punkcie leżącym w połowie odległości między ładunkami $Q_1 = 70$ nC i $Q_2 = 50$ nC, odległymi od siebie o $r = 20$ cm, umieszczonymi w nafcie o względnej przenikalności elektrycznej $\varepsilon_r = 2$.

Odp. $9 \cdot 10^3$ N/C; $1{,}1 \cdot 10^5$ V.

35. Obliczyć natężenie pola elektrycznego w odległości r od nieskończonej, nieprzewodzącej płaszczyzny naładowanej tak, że gęstość powierzchniowa jest wszędzie jednakowa i równa σ. Czy wynik zależy od odległości od płaszczyzny?

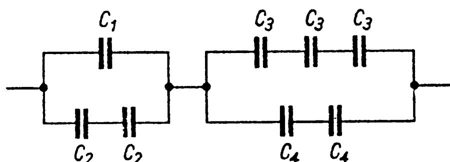
Odp. $\sigma/2\varepsilon_0\varepsilon_r$.

36. W oparciu o prawo Gaussa wykazać, że wydrążona kula o promieniu R naładowana ładunkiem Q wytwarza a) na swej powierzchni i w swoim otoczeniu zewnętrznym ($r \geqslant R$) pole elektryczne E takie samo, jakie wytwarza ładunek punktowy Q umieszczony w jej środku, b) wewnątrz kuli ($r < R$) pole o natężeniu zerowym.

37. Jaki potencjał istnieje na powierzchni cząstki α (jądro helu, $Z = 2$), jeśli przyjąć, że ma ona kształt kuli o promieniu $R = 10^{-14}$ m. Ładunek protonu $= 1{,}6 \cdot 10^{-19}$ C. (Wskazówka — por. zad. 36.)

Odp. $2{,}9 \cdot 10^5$ V.

38. Obliczyć potencjał w otoczeniu dipola (ładunki Q i $-Q$, odległość a) biorąc kolejno pod uwagę punkty: 1) na osi dipola, 2) na symetralnej dipola, 3) w odległości r_1 i r_2 odpowiednio od obu ładunków dipola.



Rys. 19.37

39. Obliczyć potencjał w środku kwadratu o boku $a = 50$ cm, jeśli w jego wierzchołkach $ABCD$ znajdują się ładunki: $Q_A = 2 \cdot 10^{-8}$ C, $Q_B = -3 \cdot 10^{-8}$ C, $Q_C = +4 \cdot 10^{-8}$ C, $Q_D = -1 \cdot 10^{-8}$ C.

Odp. 507 V.

40. W jakiej wzajemnej odległości znajdują się dwie powierzchnie ekwipotencjalne, odpowiadające różnicy potencjałów 10 V w otoczeniu nieskończonej nieprzewodzącej płyty, na której gęstość powierzchniowa ładunku $\sigma = 0{,}5 \cdot 10^{-6}$ C/m².

Odp. 0,36 mm.

41. Obliczyć pojemność zastępczą układu kondensatorów przedstawionego na rys. 19.37.

Odp. $\dfrac{(2C_1+C_2)(2C_3+3C_4)}{2(2C_3+3C_4)+6(2C_1+C_2)}$.

42. Dwa kondensatory o pojemnościach C_1 i C_2 połączono szeregowo i dołączono do zacisków baterii dającej napięcie U. Obliczyć elektryczną energię potencjalną każdego z kondensatorów.

Odp. $\dfrac{C_1C_2^2U^2}{2(C_1+C_2)}$, $\dfrac{C_2C_1^2U^2}{2(C_1+C_2)}$.

ROZDZIAŁ 20

Prąd elektryczny stały

W rozdziale poprzednim zajmowaliśmy się zagadnieniami elektrostatyki, dotyczącymi ładunków spoczywających. Tematem niniejszego rozdziału i paru następnych będą zjawiska i prawa związane z ruchem ładunków.

20.1. Natężenie prądu elektrycznego

Przez przepływ prądu elektrycznego rozumiemy ruch ładunków elektrycznych. Czynnikiem wywołującym ten ruch jest istnienie napięcia, czyli różnicy potencjałów.

W każdym zamkniętym obwodzie prądu można wyróżnić źródło (czyli tzw. część wewnętrzną obwodu) wytwarzające różnicę potencjałów między dwoma biegunami, dodatnim i ujemnym, oraz odbiorniki prądu (czyli tzw. część zewnętrzną obwodu, utworzoną z przewodników elektryczności).

Zgodnie z tradycją, za kierunek prądu w obwodzie zewnętrznym przyjmuje się kierunek od potencjału wyższego — dodatniego do niższego — ujemnego.

W czasie przepływu prądu przez przewodniki metalowe mamy do czynienia z ruchem swobodnych elektronów, a więc nośników prądu poruszających się od potencjału niższego do wyższego, czyli w kierunku przeciwnym do umownie przyjętego. W elektrolitach wchodzących w skład zewnętrznej części obwodu mamy do czynienia z ruchem jonów dodatnich do elektrody ujemnej i jonów ujemnych do elektrody dodatniej (prąd jonowy). W półprzewodnikach może występować zarówno przewodnictwo jonowe, jak i elektronowe. Podobna sytuacja istnieje w gazach.

Przez *natężenie prądu elektrycznego* (zwane też krótko *prądem elektrycznym*) rozumiemy stosunek ładunku przepływającego przez poprzeczny przekrój przewodnika do czasu przepływu:

$$I = \frac{dQ}{dt}, \tag{20.1}$$

gdzie I oznacza natężenie prądu elektrycznego, Q — ładunek elektryczny, t — czas. W przypadku prądu stałego, tj. prądu płynącego w jednym kierunku, gdy jego natężenie jest niezależne od czasu

$$I = \frac{Q}{t}. \tag{20.2}$$

Jednostką natężenia prądu elektrycznego jest amper (A). Jest to jedna z podstawowych jednostek układu SI, której definicję podaliśmy już w § 3 „Wiadomości wstępnych". Tutaj tylko przypominamy, że z równania (20.1) wynika pośrednio definicja jednostki ładunku elektrycznego, czyli 1 kulomba, gdyż $Q = \int I dt$.

O źródłach prądu stałego — ogniwach, akumulatorach, prądnicach itp. — nie będziemy tu szerzej mówili, gdyż zasada ich budowy i działania znana jest z kursu szkolnego. Podkreślimy tylko, że wielkością fizyczną, charakteryzującą źródło prądu, jest jego *siła elektromotoryczna* (*skrót SEM*). Jest to różnica potencjałów panująca na biegunach *źródła otwartego*, tj. takiego, z którego nie czerpiemy prądu. Po zamknięciu obwodu — kosztem SEM powstaje spadek potencjału wzdłuż obwodu zewnętrznego i spadek potencjału wewnątrz źródła między jego biegunami.

20.2. Prawo Ohma

Prawo Ohma, sformułowane w roku 1827 w oparciu o doświadczenia, mówi o *prostej proporcjonalności prądu I płynącego przez przewodnik do napięcia U przyłożonego na jego końcach*:

$$I = \frac{U}{R} = \frac{V_1 - V_2}{R}, \tag{20.3}$$

a więc

$$U = IR, \tag{20.4}$$

gdzie R oznacza współczynnik proporcjonalności zwany *oporem elektrycznym przewodnika*. Opór elektryczny R (zwany też *rezystancją*) wyrażany jest w omach (Ω). Z jednostek krotnych względem oma przypomnimy cztery najczęściej stosowane:

kiloom $k\Omega = 10^3\Omega$ miliom $m\Omega = 10^{-3}\Omega$

megaom $M\Omega = 10^6\Omega$ mikroom $\mu\Omega = 10^{-6}\Omega$

Opór przewodnika równa się 1 omowi, jeżeli niezmienne napięcie równe 1 woltowi istniejące na końcach przewodnika wywołuje w nim prąd o natężeniu 1 ampera:

$$1\,\Omega = \frac{1\text{V}}{1\text{A}}.$$

Odwrotność oporu elektrycznego przewodnika nosi nazwę *przewodności elektrycznej* (lub *konduktancji*). Jednostką przewodności jest *simens* (S).

$$1\,\text{S} = \frac{1\text{A}}{1\text{V}}.$$

Wynik badania zależności oporu jednorodnego, liniowego przewodnika od jego długości i od pola jego przekroju można napisać w następującej postaci:

$$R = \varrho \frac{l}{S}, \tag{20.5}$$

gdzie l oznacza długość przewodnika, S — pole jego przekroju poprzecznego, ϱ — *opór elektryczny właściwy*, zwany też *rezystywnością przewodnika*. Opór elektryczny właściwy jest to wielkość charakterystyczna dla rodzaju materiału. Opór właściwy wyraża liczbowo opór sześcianu o krawędzi 1 metra przy przepływie prądu od jednej ściany do ściany przeciwległej. Tak określony opór właściwy wyraża się w $\Omega \cdot m$*.

Odwrotność oporu właściwego przewodnika nosi nazwę *przewodności elektrycznej właściwej* (lub *konduktywności*):

$$\gamma = \frac{1}{\varrho}. \tag{20.6}$$

Jednostką konduktywności jest *simens na metr* (S/m).

W tabeli 20.1 zestawione są przykładowo wartości oporów właściwych i przewodności elektrycznych właściwych różnych materiałów.

Uwzględniając (20.5) wzór Ohma można napisać w postaci

$$I = \frac{U}{\varrho \frac{l}{S}} = \frac{1}{\varrho}\frac{U}{l}S.$$

Tabela 20.1

Opory elektryczne właściwe, przewodności elektryczne właściwe i współczynniki temperaturowe oporu elektrycznego

(temperatura około 20°C)

Rodzaj materiału	Opór elektryczny właściwy $10^{-8}\ \Omega \cdot m$	Przewodność elektryczna właściwa 10^{6} S/m	Współczynnik temperaturowy oporu $10^{-3}\ K^{-1}$
srebro	1,6	63	3,6
miedź	1,7	57	3,9
glin	2,7	37	4,0
cynk	5,9	17	3,8
żelazo	9,8	10	4,5
manganin (70% Cu, 30% Mn)	43	2,6	0,01
konstantan (35–55% Ni, 65–45% Cu, 0–20% Zn)	50	2,1	0,005
rtęć	95	1,05	0,9
wolfram	5,5	18,2	4,1
węgiel	4000	0,025	0,8

* W technice opór właściwy często jest definiowany jako wielkość wyrażająca liczbowo opór przewodnika o długości 1 metra i przekroju poprzecznym 1 mm². Wtedy opór właściwy wyraża się w $\Omega \cdot mm^2/m$. Łatwo sprawdzić, że nie ma zgodności liczbowych ϱ wynikających z obu definicji.

$$1\ \Omega \cdot m = 10^6\ \Omega \cdot mm^2/m.$$

Ale $U/l = E$ (por. wzór 19.38), a zatem

$$\frac{I}{S} = \gamma E.$$

Stosunek I/S nazywamy *gęstością prądu* i oznaczamy symbolem j. Prawo Ohma w nowej postaci sprowadza się zatem do tego, że gęstość prądu jest wprost proporcjonalna do natężenia pola elektrycznego E:

$$j = \gamma E. \tag{20.7}$$

Wyprowadzenie prawa Ohma w oparciu o klasyczną elektronową teorię przewodnictwa metali podane jest w § 20.10.

20.3. Zależność oporu elektrycznego metali od różnych czynników

Doświadczenia wykazują, że opór elektryczny przewodników metalowych zależy od takich czynników, jak stopień czystości chemicznej, temperatura, oddziaływania mechaniczne itp. Omówimy krótko niektóre z tych zależności.

Z badań zależności oporu elektrycznego metali od stopnia ich czystości wynika, że w przypadku niektórych metali drobne nawet domieszki innych pierwiastków powodują duże zmiany oporu. Tak np. stwierdzono 10-procentowe zmiany oporu miedzi przy $\frac{1}{10}$ % domieszki żelaza i $\frac{1}{100}$ % domieszki fosforu.

Opór elektryczny metali jest funkcją temperatury i na ogół ze wzrostem temperatury rośnie. W temperaturach bliskich pokojowej dla wielu metali można tę zależność traktować jako liniową:

$$R = R_0(1+\alpha t), \tag{20.8}$$

gdzie R i R_0 oznaczają odpowiednio opór przewodnika w temperaturze t i w temperaturze 0°C, α — współczynnik temperaturowy oporu elektrycznego.

Definicja współczynnika α sprowadza się do równania

$$\alpha = \frac{1}{R_0}\frac{dR}{dt}.$$

Jak widać, α przedstawia względną zmianę oporu przy zmianie temperatury o 1°C i wyraża się w K^{-1}. Wartości liczbowe współczynników temperaturowych (dla temperatur bliskich 20°C) kilku szerzej stosowanych przewodników elektrycznych uwzględniono w tab. 20.1.

O zależności oporu elektrycznego od temperatury mówiliśmy już szerzej w § 13.4 i 17.11 (termometry oporowe i nadprzewodnictwo).

Na zmianach oporu elektrycznego metalu w zależności od temperatury opiera się działanie przyrządu zwanego *bolometrem*, służącego do pomiaru natężenia promieniowania. Zasadniczą część bolometru stanowią taśmy z platyny lub niklu o grubości 0,05 μm, zgięte kilkakrotnie pod kątem prostym i pokryte sadzą dla zwiększenia absorpcji promieniowania. Ramkę bolometru ze wspomnianymi taśmami umieszcza się w osłonie próżniowej. Pod wpływem padającego promieniowania metal się ogrzewa

i zmienia się jego opór. Po uprzednim wycechowaniu ze zmian oporu wnioskuje się o natężeniu padającego promieniowania.

Niektóre stopy metali mają współczynniki temperaturowe ujemne, tzn. ich opór maleje ze wzrostem temperatury. Istnieją też stopy takie, jak np. manganin i konstantan, których opory bardzo mało zależą od temperatury. Z tym się wiąże zastosowanie tych stopów do wyrobu wzorców oporu.

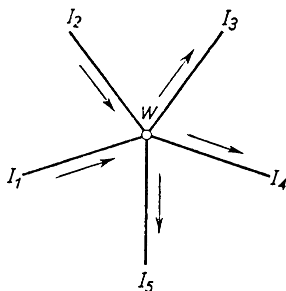
Porównanie wartości oporu właściwego metali w stanie ciekłym i stałym prowadzi do wniosku, że opór metali w stanie ciekłym jest w większości przypadków około dwa razy większy niż w stanie stałym. Wyjątek stanowi bizmut, którego opór po stopieniu maleje.

Opór przewodników metalowych zależy też od oddziaływań mechanicznych. Tak np. przy rozciąganiu drutów rośnie l, a maleje S. Oba czynniki wpływają na wzrost oporu. Na tej zależności opiera się działanie tzw. *tensometrów*. Są to przyrządy, w których wywołując zmianę długości cienkiego drutu, np. z konstantanu, uzyskuje się mierzalną zmianę oporu. Drut taki jest przyklejony do badanego elementu w ten sposób, by odkształcaniu się elementu towarzyszyło wydłużanie się drutu. Ze zmian oporu wyciąga się wnioski co do wartości odkształceń i naprężeń.

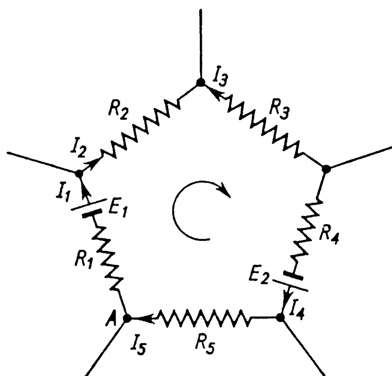
20.4. Prawa Kirchhoffa

Prawa Kirchhoffa, sformułowane w 1847 roku, ujmujemy obecnie w następującej postaci:

1. *W dowolnym punkcie W obwodu (w węźle) suma algebraiczna natężeń prądów stałych dopływających i odpływających równa się zeru.* Natężenia prądów dopływających uważamy za dodatnie, natężenia prądów odpływających za ujemne. Innymi słowy,



Rys. 20.1



Rys. 20.2

w żadnym punkcie obwodu ładunki się nie gromadzą, nigdzie też nie giną, ani nie powstają (*zasada zachowania ładunku*). Ile ładunków do węzła dopływa, tyle w tym samym czasie z niego odpływa:

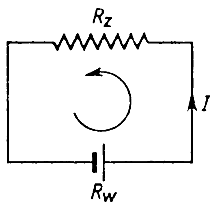
$$\sum I = 0. \tag{20.9}$$

P y t a n i e. W przypadku prądu stałego w obwodzie nie rozgałęzionym natężenie prądu I we wszystkich poprzecznych przekrojach przewodnika, niezależnie od pola ich powierzchni, ma tę samą wartość. Czy prąd elektryczny stały nie wykazuje analogii do przepływu cieczy doskonałej? Jaka wielkość charakteryzująca przepływ cieczy odpowiada natężeniu prądu?

2. *W dowolnie wydzielonej zamkniętej części obwodu elektrycznego, w tzw. oczku, suma algebraiczna wszystkich napięć elektrycznych panujących na poszczególnych elementach oczka równa się zeru.* Bierzemy tu pod uwagę wszystkie czynne siły elektromotoryczne (SEM) E, jak również wszystkie istniejące w tej części obwodu spadki potencjałów (rys. 20.2).

$$\sum U = \sum E + \sum IR = 0. \tag{20.10}$$

Przy stosowaniu wzoru (20.10) trzeba pamiętać o regule znaków, przypisującej znaki plus lub minus iloczynom IR oraz siłom elektromotorycznym źródeł prądu. Dowolny węzeł oczka (np. punkt A na rys. 20.2) przyjmujemy za punkt początkowy obiegu i w środku oczka zaznaczamy wybrany dowolnie kierunek obiegu, np. zgodnie z ruchem wskazówki zegara. Na tych odcinkach oczka, gdzie kierunek prądu jest zgodny z wybranym kierunkiem obiegu, iloczyny IR traktujemy jako dodatnie (np. $+I_1R_1$, lecz $-I_3R_3$). Siłom elektromotorycznym przypisujemy znak plus, gdy kierunek od bieguna dodatniego do ujemnego jest zgodny z wybranym kierunkiem obiegu. A zatem w odniesieniu do obwodu z rys. 20.2 wartościom E_1 i E_2 przypisujemy znak –.



Rys. 20.3

Zgodnie z omówioną regułą znaków, drugie prawo Kirchhoffa dla obwodu zaznaczonego na rys. 20.2 przyjmuje postać

$$I_1R_1+I_1R_{1w}-E_1+I_2R_2-I_3R_3+I_4R_4+I_4R_{2w}-E_2+I_5R_5 = 0,$$

gdzie R_{1w} i R_{2w} oznaczają opory wewnętrzne obu ogniw. W odniesieniu do najprostszego obwodu pojedynczego ogniwa o oporze wewnętrznym R_w (rys. 20.3), zamkniętego oporem R_z, znajdziemy

$$IR_z+IR_w-E = 0,$$

czyli

$$E = I(R_z+R_w). \tag{20.11}$$

20.5. Łączenie oporów

Rozróżniamy łączenie szeregowe i równoległe oporów.

1. Łączenie szeregowe oporów

Rozważmy obwód przedstawiony na rys. 20.4, zawierający trzy opory R_1, R_2 i R_3

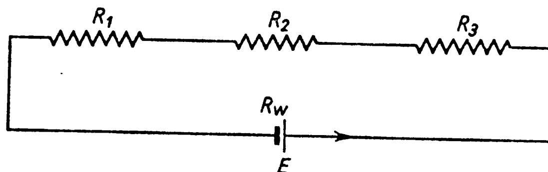
połączone szeregowo. Według II prawa Kirchhoffa (20.11)

$$E = I(R_1+R_2+R_3+R_w),$$

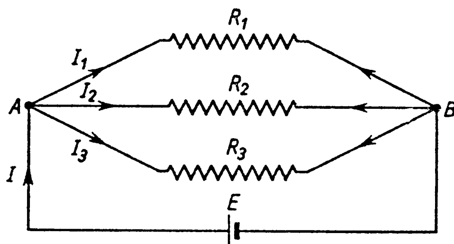
czyli całkowity opór zewnętrzny obwodu

$$R = R_1+R_2+R_3. \tag{20.12}$$

Oczywiście, wzór ten można uogólnić na dowolną liczbę połączonych szeregowo oporów. *Opór równoważny przewodników połączonych szeregowo* (zwany też oporem *zastępczym*) *równa się więc sumie oporów poszczególnych przewodników.*



Rys. 20.4



Rys. 20.5

2. Łączenie równoległe oporów

Opory połączone są równolegle, jeśli na ich końcach (A i B) (rys. 20.5) istnieje ta sama różnica potencjałów $V_A - V_B = U$. Stosując II prawo Kirchhoffa do zamkniętej części obwodu obejmującej rozgałęzienie o oporach R_1 i R_2 znajdujemy

$$R_1I_1 - R_2I_2 = 0,$$

skąd

$$\frac{I_1}{I_2} = \frac{R_2}{R_1}.$$

Analogicznie dla zamkniętej części obwodu zawierającej opory R_2 i R_3 znajdziemy

$$\frac{I_2}{I_3} = \frac{R_3}{R_2}.$$

Prądy w poszczególnych gałęziach są więc odwrotnie proporcjonalne do oporów tych gałęzi.

Aby wyznaczyć opór R równoważny oporowi n przewodników połączonych równolegle, zastosujemy prawo Ohma kolejno do poszczególnych gałęzi:

$$I_1 = \frac{U}{R_1},$$

$$I_2 = \frac{U}{R_2},$$

$$I_3 = \frac{U}{R_3},$$

$$\cdots\cdots\cdots$$

$$I_n = \frac{U}{R_n}.$$

Sumując stronami otrzymujemy

$$I_1+I_2+I_3+\ldots+I_n = U\left(\frac{1}{R_1}+\frac{1}{R_2}+\frac{1}{R_3}+\ldots+\frac{1}{R_n}\right).$$

Uwzględniając I prawo Kirchhoffa,

$$I_1+I_2+I_3+\ldots+I_n = I,$$

i porównując z ogólnym wzorem Ohma,

$$I = \frac{U}{R},$$

znajdujemy, że

$$\frac{1}{R} = \frac{1}{R_1}+\frac{1}{R_2}+\frac{1}{R_3}+\ldots+\frac{1}{R_n}. \qquad (20.13)$$

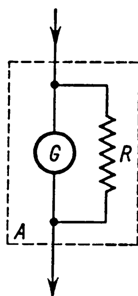
A zatem *odwrotność oporu równoważnego oporowi n przewodników połączonych równolegle równa się sumie odwrotności oporów składowych.*

Z otrzymanych zależności dla łączenia szeregowego i równoległego przewodników wynika, że łączenie szeregowe zwiększa opór równoważny (a więc zmniejsza natężenie prądu przy danym napięciu źródła), natomiast łączenie równoległe opór ten zmniejsza. Łatwo sprawdzić, że opór równoważny n jednakowych przewodników o oporze R' połączonych szeregowo wynosi nR', a połączonych równolegle R'/n.

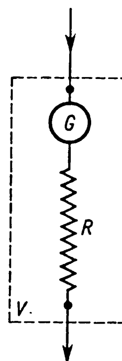
Warto tu wspomnieć o niektórych zastosowaniach łączenia szeregowego i równoległego oporów. Wiemy, że zasada budowy amperomierzy i woltomierzy jest podobna. Amperomierz powinien jednak mieć opór mały, aby jego szeregowe włączenie do obwodu możliwie jak najmniej wpływało na zmierzoną wartość natężenia prądu. Zmniejszenie oporu można uzyskać włączając równolegle do głównej części G przyrządu opornik o małym oporze R, zwany *bocznikiem* (rys. 20.6). Wtedy całości urządzenia zapewniamy mały opór równoważny. Odwrotnie — woltomierz powinien mieć opór możliwie duży, aby jego równoległe włączenie między dwa punkty obwodu jak najmniej wpływało na prąd w tym obwodzie, a tym samym i na mierzoną wartość napięcia między wybranymi punktami. Duży opór równoważny całości urządzenia uzyskujemy

włączając szeregowo do głównej części przyrządu G opornik o dużym oporze R (rys. 20.7), zwany *opornikiem szeregowym.*

Amperomierze i woltomierze są często zaopatrywane w komplety dodatkowych oporników. W przypadku amperomierzy są to wymienne boczniki, w przypadku woltomierzy — wymienne oporniki szeregowe. Jedne i drugie mogą być w łatwy sposób dołączane do właściwego przyrządu pomiarowego. Zastosowanie ich zmienia zakres skali przyrządu. Rozważmy to na przykładzie amperomierza, pozostawiając przypadek woltomierza do samodzielnego rozpatrzenia (por. zad. 25).



Rys. 20.6



Rys. 20.7

Niech opór amperomierza wynosi R_0, opór bocznika R_b. Chcemy ustalić warunek, jaki musi spełniać R_b, by prąd płynący przez R_0 stanowił n-tą część całkowitego prądu I. Jeśli przez R_0 płynie prąd I/n, to przez R_b — zgodnie z I prawem Kirchhoffa — powinien płynąć prąd

$$\frac{(n-1)}{n} I.$$

Ale między prądami w odnogach i ich oporami istnieje zależność odwrotnie proporcjonalna, a zatem

$$\frac{R_0}{R_b} = \frac{(n-1)In}{nI},$$

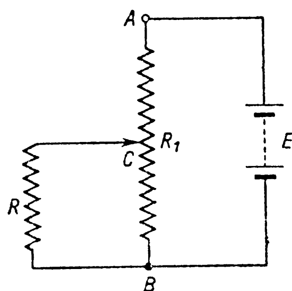
skąd

$$R_b = \frac{R_0}{(n-1)}.$$

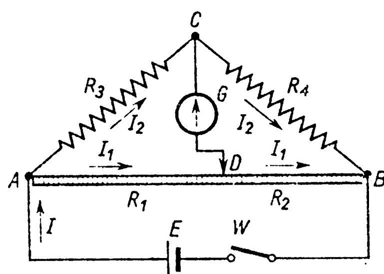
Jeżeli więc chcemy, by przez przyrząd płynął prąd stanowiący np. dziesiątą ($n = 10$) część prądu całkowitego, musimy zastosować bocznik o oporze równym dziewiątej części ($n-1 = 9$) oporu miernika. Wtedy odczytując na skali np. 0,02 A wnioskujemy, że prąd płynący w obwodzie głównym ma natężenie 0,2 A. Jeżeli skala przyrządu sięga do maksymalnej wartości 0,1 A, to przy zastosowaniu omawianego bocznika można mierzyć prądy do 1 A.

Z łączeniem oporów mamy też do czynienia przy stosowaniu tzw. *potencjometrów,* urządzeń służących do regulacji napięcia na końcach obwodu roboczego o oporze R.

Zasada działania potencjometru przedstawiona jest na rys. 20.8. Zaciski A i B opornika R_1 łączymy ze źródłem prądu stałego, uzyskując w ten sposób ciągły spadek potencjału od wartości V_A w punkcie A do wartości V_B w punkcie B. Do zacisku w punkcie B dołączony jest na stałe jeden z końców obwodu roboczego, podczas gdy drugi — w postaci styku ślizgowego — można dowolnie przesuwać wzdłuż opornika R_1. Zmieniając położenie styku C uzyskujemy na końcach obwodu roboczego napięcie zmieniające się od zera (gdy styk C przypada w punkcie B) do maksymalnej wartości $V_A - V_B$ (gdy styk przypada w punkcie A).



Rys. 20.8



Rys. 20.9

Jako przykład zastosowania praw Ohma i Kirchhoffa do obwodu elektrycznego rozpatrzmy metodę pomiaru oporu elektrycznego zwaną metodą mostka Wheatstone'a (rys. 20.9). Obwód mostka składa się ze źródła napięcia E, drutu oporowego o stałym przekroju rozciągniętego na tle skali liniowej od A do B, opornika pomocniczego o oporze znanym R_3, opornika o oporze badanym R_4, galwanometru zerowego, łącznika jednobiegunowego (klucza) W i przewodów łącznikowych.

W punktach A i B mamy rozgałęzienia obwodu. Prąd o natężeniu I płynący od źródła dzieli się na prąd I_1, płynący przez drut oporowy AB, i na prąd I_2, płynący przez oporniki R_3 i R_4. Zgodnie z pierwszym prawem Kirchhoffa

$$I = I_1 + I_2.$$

Zarówno wzdłuż jednej, jak i drugiej drogi prądu mamy ciągły spadek potencjału od V_A do V_B. Oznaczmy potencjał w punkcie C przez V_C. Na drucie oporowym AB można też znaleźć taki punkt D, w którym potencjał

$$V_D = V_C.$$

Doświadczalnie znajdujemy położenie tego punktu przesuwając ruchomy styk wzdłuż drutu AB do takiego położenia, aby wskazówka galwanometru zerowego G włączonego między C i D nie odchylała się od zera. Takie zachowanie się wskazówki świadczy o tym, że między punktami C i D nie ma różnicy potencjałów.

Wypisujemy dla takich warunków prawo Ohma kolejno dla poszczególnych części obwodu:

$$V_A - V_C = I_2 R_3,$$

$$V_A - V_D = I_1 R_1,$$
$$V_C - V_B = I_2 R_4,$$
$$V_D - V_B = I_1 R_2.$$

Z pierwszych dwóch równań otrzymujemy

$$I_2 R_3 = I_1 R_1,$$

z następnych

$$I_2 R_4 = I_1 R_2,$$

skąd

$$\frac{R_3}{R_4} = \frac{R_1}{R_2}.$$

Pamiętając o zależności (20.5) otrzymujemy

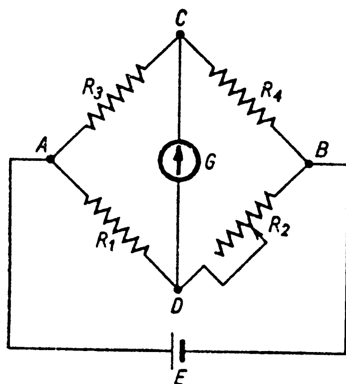
$$\frac{R_3}{R_4} = \frac{\varrho l_1 S}{\varrho l_2 S},$$

skąd

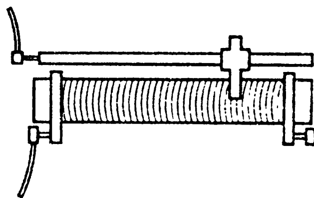
$$R_4 = \frac{R_3 l_2}{l_1},$$

gdzie l_1 i l_2 są długościami *AD* i *DB* drutu oporowego, odczytanymi bezpośrednio na podziałce.

Pomiar oporu metodą mostka Wheatstone'a jest pomiarem względnym, wymaga bowiem znajomości oporu pomocniczego R_3. W zasadzie mierząc długości l_1 i l_2 otrzymujemy dane potrzebne do ustalenia stosunku R_4 do R_3.



Rys. 20.10

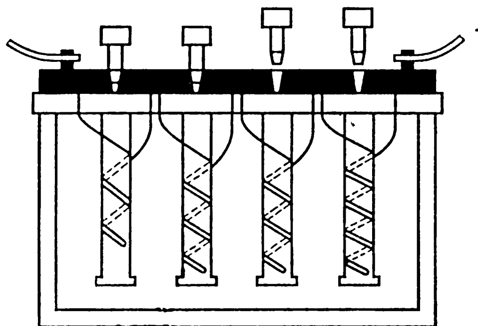

Rys. 20.11

Oczywiście, metoda mostka Wheatstone'a może być stosowana przy użyciu nieco innych układów. Na rys. 20.10 przedstawiony jest schemat mostka z regulowanym jednym oporem (R_2). Przez zmianę oporu R_2 doprowadza się mostek do „zrównoważenia", tzn. do takiego stanu, że przez galwanometr *G* nie płynie prąd. Wtedy — jak poprzednio — obowiązuje zależność

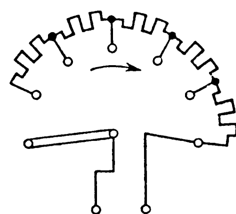
$$\frac{R_1}{R_2} = \frac{R_3}{R_4},$$

z której można wyznaczyć dowolny opór przy znanych trzech pozostałych.

Opór obwodów elektrycznych można doprowadzić do żądanej wartości przez stosowanie różnych typów oporników. Rysunki 20.11, 20.12 i 20.13 przedstawiają kolejno schematy oporników: suwakowego, wtyczkowego i korbowego. Zasada ich działania znana jest ze szkoły średniej.



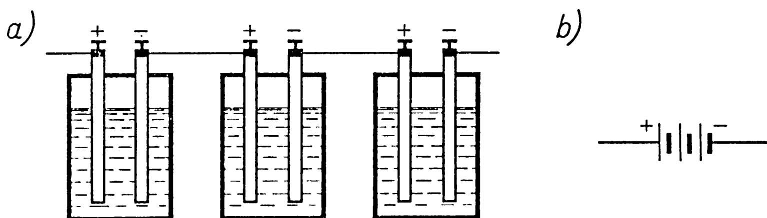
Rys. 20.12



Rys. 20.13

20.6. Łączenie źródeł napięcia prądu stałego

Omówienie prawa Ohma byłoby niepełne, gdybyśmy poprzestali na podaniu wzorów (20.4) dla obwodu otwartego i (20.11) dla najprostszego obwodu zamkniętego. Musimy jeszcze uwzględnić przypadki, gdy źródła napięcia prądu stałego, np. ogniwa, akumulatory, są połączone szeregowo lub równolegle w baterie.



Rys. 20.14

1. Łączenie szeregowe ogniw (rys. 20.14a i b) sprowadza się do łączenia bieguna ujemnego jednego ogniwa z biegunem dodatnim następnego ogniwa. Przy takim sposobie łączenia sumują się siły elektromotoryczne i opory wewnętrzne poszczególnych ogniw. A zatem przy szeregowym połączeniu n jednakowych ogniw między biegunem dodatnim pierwszego i biegunem ujemnym ogniwa n-tego istnieje SEM równa nE. Opór wewnętrzny takiej baterii ogniw równa się nR_w, gdzie R_w jest oporem wewnętrznym pojedynczego ogniwa. Po zamknięciu obwodu opornikiem zewnętrznym o oporze R_z otrzymujemy prąd I równy

$$I = \frac{nE}{R_z + nR_w}. \qquad (20.14)$$

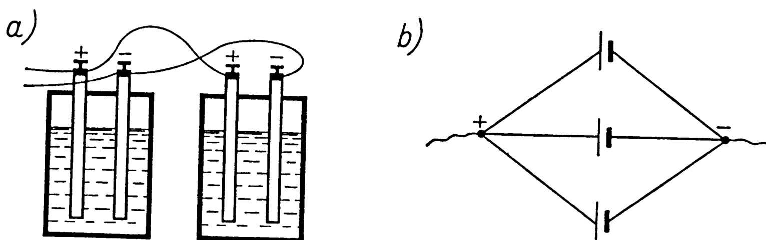
Szeregowy sposób łączenia ogniw sprzyja uzyskaniu większego prądu w przypadku, gdy opór wewnętrzny R_w jest mały w porównaniu z oporem zewnętrznym. Z przekształconego równania (20.14):

$$I=\frac{E}{\frac{R_z}{n}+R_w}$$

wynika, że gdy $R_w \ll R_z$, to prąd w wyniku szeregowego połączenia n ogniw rośnie w przybliżeniu proporcjonalnie do n:

$$I\approx\frac{nE}{R_z}.$$

(Bardzo mały opór R_w mają akumulatory, zwłaszcza o dużych rozmiarach.)



Rys. 20.15

2. Łączenie równoległe jednakowych ogniw (rys. 20.15a i b), sprowadzające się do łączenia ze sobą wszystkich biegunów tego samego znaku, nie ma wpływu na SEM baterii. Bateria n jednakowych ogniw połączonych równolegle ma taką samą SEM, równą E, jak pojedyncze ogniwo. Zmienia się natomiast opór wewnętrzny z wartości R_w na R_w/n, czyli

$$I=\frac{E}{R_z+R_w/n}.$$

Jak widać, ten sposób łączenia ogniw n-krotnie powiększa prąd w przypadku, gdy opór zewnętrzny R_z jest tak mały, że można go zaniedbać w porównaniu z oporem wewnętrznym baterii.

Należy podkreślić następującą ważną zaletę łączenia równoległego źródeł. Każde źródło ma określoną górną granicę prądu I_{gr}, który można z niego czerpać. Przekroczenie tej granicy może doprowadzić do uszkodzenia źródła. Stosując połączenie równoległe n jednakowych źródeł można — bez wywołania ich uszkodzeń — uzyskać w obwodzie zewnętrznym prąd nI_{gr} przy stałym napięciu na oporze R_z.

3. Łączenie szeregowo-równoległe ogniw (rys. 20.16). Załóżmy, że mamy m grup zawierających po n jednakowych, szeregowo połączonych ogniw. Grupy są połączone równolegle. W tym przypadku czynna siła elektromotoryczna równa się nE, a równoważny opór wewnętrzny wynosi nR_w/m. Natężenie prądu w obwodzie wynosi

$$I=\frac{nE}{R_z+nR_w/m}.$$

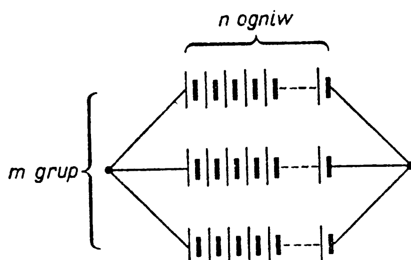
Dzieląc licznik i mianownik tego wyrażenia przez n przejdziemy do postaci

$$I = \frac{E}{\frac{R_z}{n} + \frac{R_w}{m}}. \tag{20.15}$$

To wyrażenie, jak łatwo sprawdzić przez różniczkowanie, osiąga maksimum wtedy, gdy

$$\frac{R_z}{n} = \frac{R_w}{m}, \quad \text{czyli} \quad R_z = \frac{nR_w}{m},$$

tzn. wtedy, gdy opór zewnętrzny obwodu równa się oporowi wewnętrznemu utworzonej baterii.



Rys. 20.16

20.7. Moc prądu elektrycznego

Aby określić pracę prądu elektrycznego i jego moc, opieramy się na wzorze na pracę przenoszenia ładunku dQ między punktami o różnicy potencjałów U:

$$dW = UdQ.$$

Uwzględniając definicję natężenia prądu (20.1) otrzymujemy

$$dW = UIdt,$$

$$W = \int UIdt. \tag{20.16a}$$

Przy założeniu, że $U = \text{const}$ i $I = \text{const}$, praca prądu wyraża się wzorem

$$W = UIt. \tag{20.16b}$$

Oczywiście, to samo wyrażenie przedstawia energię cieplną (w dżulach) wydzieloną w obwodzie podczas przepływu prądu.

Od wyrażenia na pracę przechodzimy do wzoru na moc prądu elektrycznego:

$$P = \frac{dW}{dt} = UI. \tag{20.17}$$

Podstawienie do tego wzoru napięcia w woltach i prądu w amperach daje moc w watach.

Uwzględniając prawo Ohma można wzór (20.16b) napisać w postaci

$$W = I^2Rt \tag{20.18}$$

lub

$$W = \frac{U^2}{R}t. \tag{20.19}$$

Można też wzór (20.16b) przekształcić wprowadzając gęstość prądu j i uwzględniając równanie (20.7):

$$W = UjSt = U\gamma ESt.$$

Po uwzględnieniu zależności $E = U/l$ otrzymujemy

$$W = \gamma E^2 lSt.$$

Iloczyn lS wyraża objętość przewodnika, w którym wydziela się energia. Ilość energii wydzielona w jednostce czasu w jednostce objętości przewodnika wynosi

$$\frac{W}{tV} = \gamma E^2 = jE, \tag{20.20}$$

jest więc proporcjonalna do gęstości prądu j i do natężenia pola elektrycznego E.

Ze wzoru (20.18) wynika, że przy niezmiennym oporze ilość wydzielanej energii cieplnej jest proporcjonalna do kwadratu natężenia prądu. Duże prądy powodują wydzielanie dużych ilości ciepła, prowadzące nawet czasami do stopienia części instalacji.

Pytanie. Dlaczego drucik żarówki w czasie przepływu prądu żarzy się, a przewód miedziany doprowadzający prąd pozostaje zimny? Dlaczego miejsca złego kontaktu silnie się grzeją?

Ciepło wydzielane przez prąd elektryczny jest wykorzystywane w licznych urządzeniach, jak wszelkiego rodzaju grzejniki, piece elektryczne itp. W piecach elektrycznych można topić nawet trudno topliwe metale, stosując czy to metodę łuku elektrycznego (napięcia rzędu kilkudziesięciu woltów, prądy rzędu kilkudziesięciu amperów), czy też metodą kontaktową (prądy rzędu tysięcy amperów).

Wywołując przepływ dużych prądów przez cienkie druciki można powodować ich rozpylanie przy oślepiającym świetle (temperatury do 20 000°C).

20.8. Prąd elektryczny w elektrolitach

20.8.1. *Prawa Faradaya.* W grupie dobrych przewodników elektryczności umieściliśmy przede wszystkim metale, następnie tzw. elektrolity. *Przewodnictwo metaliczne*, związane z ruchem swobodnych elektronów, różni się dość zasadniczo od *przewodnictwa elektrolitycznego*. W pierwszym przypadku przepływowi prądu nie towarzyszą na ogół żadne chemiczne zmiany przewodnika. Zmiany takie są natomiast charakterystyczne dla przepływu prądu przez elektrolity.

Najpospolitszymi *elektrolitami* są roztwory (przede wszystkim wodne) kwasów, zasad i soli. Liczne roztopione sole również przewodzą prąd elektryczny. W wyniku przepływu prądu na elektrodzie ujemnej — *katodzie* — wydzielają się takie substancje, jak wodór, metale oraz grupy takie jak NH_4. Na elektrodzie dodatniej — *anodzie* — wydzielają się: tlen, reszty kwasowe, grupa OH. Wydzielanie się substancji w wyniku

przepływu prądu nazywamy *elektrolizą*. Naczynie, w którym odbywa się elektroliza, nosi nazwę *woltametru*.

Badaniem elektrolizy zajmował się Faraday, który wyniki swych doświadczeń ujął w dwa prawa, zwane dziś *prawami Faradaya*:

1. *Masy produktów elektrolizy wydzielone na elektrodach są proporcjonalne do natężenia prądu i czasu jego przepływu, czyli do ładunku przepływającego przez elektrolit*:

$$m = kIt,$$
$$m = kQ. \tag{20.21}$$

Współczynnik proporcjonalności k nosi nazwę *równoważnika elektrochemicznego*. Wyraża on liczbowo masę produktu elektrolizy wydzieloną na elektrodzie przez prąd o natężeniu 1 ampera w ciągu 1 sekundy, czyli podczas przepływu przez elektrolit ładunku 1 kulomba. W tabeli 20.2 podane są wartości równoważników elektrochemicznych dla kilku substancji.

Tabela 20.2

Równoważniki elektrochemiczne

Rodzaj substancji	k 10^{-7} kg/A · s
cynk	3,388
glin	0,933
miedź	3,294
srebro	11,18
wodór	0,104

2. *Masy produktów elektrolizy wydzielone na elektrodach różnych woltametrów podczas przepływu prądu o tym samym natężeniu i w tym samym czasie są proporcjonalne do gramorównoważników danych substancji*:

$$m_1 : m_2 : m_3 \ldots = R_1 : R_2 : R_3 \ldots \tag{20.22}$$

Gramorównoważnik danej substancji jest to stosunek masy gramoatomu pierwiastka lub gramocząsteczki (mola) M związku chemicznego do wartościowości

$$R = \frac{M}{w}.$$

Oba prawa Faradaya można ująć w jedno w następujący sposób. Masy wydzielone w różnych woltametrach w jednakowych czasach t przez jednakowe prądy I można wyrazić jako

$$m_1 = k_1 It,$$
$$m_2 = k_2 It \text{ itd.}$$

Dzieląc stronami znajdujemy

$$m_1 : m_2 : m_3 \ldots = k_1 : k_2 : k_3 \ldots$$

Porównując to wyrażenie z (20.22) otrzymamy

$$R_1:R_2:R_3\,... = k_1:k_2:k_3\,...,$$

czyli *stosunek gramorównoważników równa się stosunkowi równoważników elektrochemicznych danych substancji.*

Z ostatniej zależności wynika, że

$$\frac{R_1}{k_1}=\frac{R_2}{k_2}=\frac{R_3}{k_3}=\,...\,=\mathrm{const}.$$

Stałą wartość stosunku gramorównoważnika do równoważnika elektrochemicznego danej substancji nazywamy *stałą Faradaya* i oznaczamy symbolem F:

$$F=\frac{R}{k}.$$

Sens fizyczny stałej Faradaya łatwo jest ustalić. Skoro wydzielenie w czasie elektrolizy masy równej równoważnikowi elektrochemicznemu k wymaga przepływu ładunku jednego kulomba, to wydzielenie masy równej gramorównoważnikowi R wymaga przepływu ładunku równego stosunkowi R/k, czyli właśnie równego stałej Faradaya F. Wartość liczbową F obliczymy wykorzystując dane liczbowe, np. srebra:

$$F=\frac{107{,}88\ \mathrm{g}}{1{,}118\cdot 10^{-3}\ \mathrm{g\cdot C^{-1}}}=96\,487\ \mathrm{C}.$$

Taki jest zatem ładunek wydzielający w czasie elektrolizy gramorównoważnik dowolnej substancji (a więc np. 1,008 g wodoru, 35,45 g chloru, 62 g NO_3, 8 g tlenu, 48,03 g SO_4 itp.).

Korzystając z wyrażenia definiującego stałą Faradaya można napisać

$$k=\frac{R}{F}=\frac{M}{wF}$$

i można nadać prawu Faradaya ostateczną postać:

$$m=\frac{M}{wF}It. \tag{20.23}$$

20.8.2. *Mechanizm elektrolizy.* Przepływ prądu elektrycznego przez elektrolit związany jest z ruchem jonów. Jony powstają w ten sposób, że obojętne atomy lub grupy atomowe tracą lub przyłączają do siebie jeden lub więcej elektronów. Liczba utraconych lub przyłączonych elektronów decyduje o *wartościowości jonu.*

Na wynikach badań elektrolizy można oprzeć wyznaczanie ładunku elementarnego i ugruntowanie poglądu o atomistycznej budowie elektryczności. Wiemy już, że przeniesienie ładunku F przez elektrolit powoduje wydzielenie jednego gramorównoważnika dowolnej substancji, co w przypadku pierwiastka jednowartościowego sprowadza się

do wydzielenia jednego gramoatomu. Poza tym wiemy, że w jednym gramoatomie jest tyle atomów, a więc i jonów, ile wynosi liczba Avogadra — N_A. A zatem każdy jon przenosi ładunek e:

$$e=\frac{F}{N_A}=\frac{96\,487}{6{,}02\cdot 10^{23}}=1{,}6\cdot 10^{-19}\ \text{C}.$$

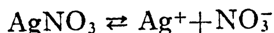
Jest to wartość ładunku elementarnego, całkowicie zgodna z wynikami otrzymanymi innymi metodami (por. § 19.20). Jony wielowartościowe mają ładunki będące całkowitymi wielokrotnościami ładunku elementarnego.

Wróćmy do procesu wytwarzania się jonów w elektrolicie. Jeśli założymy, że czynnikiem wiążącym w cząsteczkach związków chemicznych jest przyciąganie elektryczne jonów różnoimiennych, wchodzących w ich skład, to łatwo zrozumiemy, że wprowadzenie takiej cząsteczki do rozpuszczalnika o dużej przenikalności elektrycznej zmniejszy siłę wzajemnego przyciągania między jonami. Może temu towarzyszyć nawet rozpad cząsteczki na jony, czyli tzw. *dysocjacja elektrolityczna*. Jeśli np. w wodzie ($\varepsilon_r=80$) rozpuścimy jakąś sól, kwas lub zasadę, to część cząsteczek ulegnie dysocjacji. Stosunek liczby cząsteczek zdysocjowanych do ogólnej liczby cząsteczek wprowadzonych do rozpuszczalnika nazywamy *stopniem dysocjacji roztworu*.

Dla przykładu podajemy równania dysocjacji kilku ogólnie znanych soli:



$$NaCl \rightleftarrows Na^{+}+Cl^{-},$$



$$AgNO_3 \rightleftarrows Ag^{+}+NO_3^{-}$$



$$Na_2SO_4 \rightleftarrows 2Na^{+}+SO_4^{--}\ \text{itp.}$$

Wyjaśnienie elektrolizy na podstawie dysocjacji elektrolitycznej podał Arrhenius. Jest to teoria uproszczona, gdyż Arrhenius nie brał pod uwagę wzajemnego oddziaływania jonów. Dziś wiemy, że teorię Arrheniusa można stosować tylko do tzw. *elektrolitów słabych*, takich jak np. większość kwasów organicznych. *Elektrolity mocne*, do których zaliczamy wiele związków nieorganicznych, dysocjują w 100% i ich własności wynikają z wzajemnego silnego oddziaływania między jonem i jego „jonowym" otoczeniem.

Zgodnie z założeniami fizyki cząsteczkowej, jony powstałe w wyniku dysocjacji są w ciągłym chaotycznym ruchu. Możliwe przy tym jest spotkanie jonów różnoimiennych i ponowne wytworzenie obojętnej cząsteczki (*proces rekombinacji*). W tym samym czasie inna cząsteczka może ulec rozpadowi. W roztworze stale rozwijają się oba procesy: proces dysocjacji i proces rekombinacji.

Wprowadzenie do roztworu kwasu, zasady lub soli płytek metalowych połączonych z biegunami źródła napięcia wywołuje uporządkowanie ruchu jonów. Jony dodatnie (*kationy*) dążą do elektrody ujemnej, jony ujemne (*aniony*) — do elektrody dodatniej. Po dojściu do odpowiednich elektrod ładunki jonów zostają zobojętnione: jony dodatnie dołączają elektrony z katody, jony ujemne — oddają swe nadmiarowe elektrony anodzie. Czasem produkty elektrolizy działają chemicznie czy to na rozpuszczalnik, czy na materiał elektrody, wchodząc w tzw. *reakcje wtórne*. Tak na przykład podczas

elektrolizy NaCl sód działa na wodę powodując wydzielenie gazowego wodoru oraz ważnego dla przemysłu NaOH. Na drugiej elektrodzie wydziela się gazowy Cl_2.

Podczas elektrolizy Na_2SO_4 nietrwały SO_4 przechodzi z wydzieleniem tlenu w SO_3, który łącząc się z wodą daje H_2SO_4. Na drugiej elektrodzie zachodzi analogiczna do poprzednio omówionej reakcja sodu z wodą. Interesujący przebieg ma elektroliza słabego kwasu tlenowego, np. kwasu borowego, w obecności anody glinowej. W wyniku elektrolizy anoda pokrywa się cienką warstwą nierozpuszczalnego w wodzie i nieprzewodzącego tlenku Al_2O_3. Ta cienka warstwa może służyć jako ochrona aluminiowych części samochodowych lub samolotowych przed korozją. Układ złożony z płytki aluminiowej (anoda), cienkiej warstwy nieprzewodzącego tlenku i warstwy przewodzącego elektrolitu stanowi kondensator elektrolityczny o dużej pojemności dzięki bardzo małej grubości warstwy tlenku. (Można osiągać w ten sposób pojemność rzędu kilkunastu pF na 1 cm^2 powierzchni okładki. Wielkie kondensatory elektrolityczne mają pojemności do 1 mF, a nawet więcej.)

Z zastosowań elektrolizy warto wspomnieć znane ze szkoły średniej 1) galwanostegię, 2) galwanoplastykę, 3) otrzymywanie czystych pierwiastków z ich związków. Przeliczając w skali ogólnoświatowej ustalono, że na pokrywanie elektrolityczne metali mniej szlachetnych bardziej szlachetnymi, a więc na takie procesy, jak srebrzenie, niklowanie, chromowanie itp. wykorzystuje się stale — w każdej sekundzie — prąd o natężeniu rzędu 10 milionów amperów.

Metoda elektrolitycznego otrzymywania czystych metali dostarcza np. czystą elektrolityczną miedź o stopniu czystości 99,99%. Ta sama metoda jest metodą podstawową dla otrzymywania glinu (stosuje się roztopiony Al_2O_3 i Na_3AlF_6 — kriolit). Otrzymywanie czystych metali na drodze elektrolitycznej związane jest ze stałym zużyciem w skali ogólnoświatowej prądu o natężeniu około 100 milionów amperów.

Na podstawie elektrolizy można przeprowadzić chemiczną analizę soli. Okazuje się bowiem, że elektroliza każdej soli rozpoczyna się od pewnej określonej różnicy potencjałów między elektrodami, zwanej potencjałem rozkładu danej soli. Można więc przez stopniowe zwiększanie różnicy potencjałów wywoływać kolejne wydzielanie poszczególnych składników z mieszaniny soli.

Na zakończenie wzmianka o innym jeszcze zastosowaniu elektrolizy, a mianowicie o elektrolitycznym polerowaniu metali. Wygładzoną mechanicznie płytę metaliczną umieszcza się jako anodę w odpowiednim roztworze. Podczas przepływu prądu przede wszystkim wygładzają się mikroskopijne wypukłości powierzchni anody.

20.8.3. *Przewodność elektrolityczna. Ruchliwość jonów.* Powiedzieliśmy, że pole elektryczne wywołuje uporządkowany ruch jonów dodatnich i ujemnych do odpowiednich elektrod. Zachowanie się różnych jonów charakteryzuje ich ruchliwość. *Ruchliwość* (u^+ lub u^-) jonu wyraża liczbowo prędkość jonu osiąganą w polu o natężeniu jednostkowym. W tabeli 20.3 podane są ruchliwości $\left(\text{w } \frac{\text{m/s}}{\text{V/m}}\right)$ różnych jonów w roztworze wodnym w temperaturze 18°C. Wzrost temperatury silnie wpływa na ruchliwość jonów. Przyrostowi temperatury o 1°C towarzyszy wzrost ruchliwości o około 2%.

Tabela 20.3

Ruchliwości jonów w roztworach wodnych w temperaturze 18°C

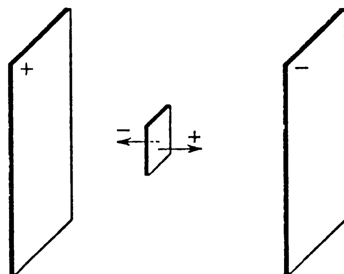
Rodzaj jonu	Ruchliwość $10^{-8}\frac{m/s}{V/m}$
Na^{+}	4,35
Ag^{+}	5,6
Zn^{++}	4,8
Cl^{-}	6,5
SO_4^{--}	6,8
OH^{-}	17,4

Przewodność właściwą elektrolitów słabych można wyrazić za pomocą ruchliwości jonów. Załóżmy, że badany roztwór jest bardzo rozcieńczony. Innymi słowy, a) odrzucamy prawdopodobieństwo łączenia się jonów w cząsteczki, b) nie bierzemy pod uwagę wzajemnego oddziaływania jonów.

Niech n oznacza liczbę jonów dodatnich i ujemnych w jednostce objętości, w — ich wartościowość. Ładunek pojedynczego jonu wynosi zatem we. Niech natężenie E pola elektrycznego między elektrodami będzie niezbyt duże. W tych warunkach możemy zakładać, że prędkości jonów v^+ i v^- są proporcjonalne do natężenia pola:

$$v^+ = u^+E,$$

$$v^- = u^-E.$$



Rys. 20.17

Przez jednostkę powierzchni pomyślanej między anodą i katodą (rys. 20.17), prostopadłej do linii sił pola, w jednostce czasu zostanie przeniesiony ładunek ujemny Q_- do anody:

$$Q_- = nu^-Ewe$$

oraz ładunek dodatni Q_+ do katody:

$$Q_+ = nu^+Ewe.$$

Przepływ ładunków ujemnych do anody jest równoważny przepływowi ładunków do-

datnich w stronę przeciwną. A zatem gęstość prądu j, czyli ładunek przepływający w jednostce czasu przez jednostkę powierzchni, wynosi

$$j = nwe(u^+ + u^-)E.$$

Innymi słowy, istnieje prosta proporcjonalność gęstości prądu do natężenia pola. Ale w § 20.2 wykazaliśmy, że prawo Ohma po przekształceniu przyjmuje analogiczną postać:

$$j = \gamma E,$$

gdzie γ jest przewodnością elektryczną właściwą. Stwierdzamy zatem, że *do elektrolitów słabych, rozcieńczonych, przy niezbyt silnych polach stosuje się prawo Ohma.* Z porównania wzorów wynika, że przewodność właściwa elektrolitu jest proporcjonalna do stężenia jonów w roztworze i do sumy ruchliwości jonów dodatnich i ujemnych:

$$\gamma_{\text{elektrolitu}} = nwe(u^+ + u^-). \tag{20.24}$$

Jest ona 10^5–10^6 razy mniejsza niż w metalach i rośnie ze wzrostem temperatury. *Roztwory elektrolitów silnych wykazują duże odstępstwa od prawa Ohma.*

20.9. Prąd elektryczny w gazach

20.9.1. *Jonizacja gazów.* Gazy są na ogół złymi przewodnikami elektryczności, w pewnych warunkach jednak — a mianowicie po zjonizowaniu — stają się przewodnikami. Jonizacja gazów ma charakter inny niż jonizacja w elektrolitach. Obojętna cząsteczka lub atom gazu traci słabiej związany elektron i przekształca się w jon dodatni. Niektóre z tych oderwanych elektronów dołączają się do obojętnych cząsteczek lub atomów, tworząc jony ujemne. Tak np. powstają jony ujemne H^-, O^-, OH^-, NO_2^-, Cl^- itd. W przypadku azotu i gazów szlachetnych nie obserwuje się dołączania elektronów do obojętnych atomów.

Oddzielenie elektronu od atomu wymaga pewnej pracy, wymaga dostarczenia *energii jonizacji**.

Zamiast o energii jonizacji często mówi się o *potencjale jonizacji*, rozumiejąc przez to różnicę potencjałów, jaką musiałby pokonać elektron w polu elektrycznym, aby związana z tym energia równała się energii jonizacji**. Potencjały jonizacji dla kilku jonów podane są w tab. 20.4.

Do *czynników jonizujących* gazy zaliczamy m. in.: 1) promieniowanie nadfioletowe, 2) promieniowanie Röntgena, 3) promieniowanie ciał promieniotwórczych (α, β, γ), 4) reakcje chemiczne, 5) wysoką temperaturę, 6) promieniowanie kosmiczne.

---

* Energię tę podajemy zazwyczaj w tzw. elektronowoltach (eV). *Elektronowolt* jest to praca związana z przeniesieniem elektronu (czyli ładunku elementarnego o wartości $e = 1{,}601 \cdot 10^{-19}$ C) między dwoma punktami pola o różnicy potencjałów 1 wolta. W przeliczeniu na dżule praca ta wynosi

$$1\ \text{eV} = 1{,}601 \cdot 10^{-19}\ \text{J}.$$

** Potencjały jonizacji wyrażone w woltach pokrywają się liczbowo z energiami jonizacji wyrażonymi w elektronowoltach.

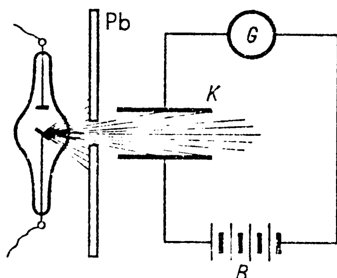
Tabela 20.4

Potencjały jonizacji

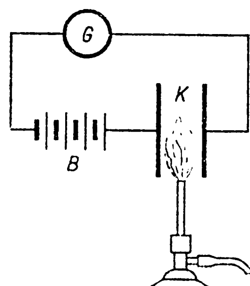
Rodzaj przejścia	Potencjał jonizacji V
$H \rightarrow H^+$	13,5
$O \rightarrow O^+$	13,5
$O_2 \rightarrow O_2^+$	12,5
$He \rightarrow He^+$	24,5

Warto zapamiętać, że samo podniesienie temperatury prowadzi do jonizacji podczas zderzeń cząsteczek w ich ruchach cieplnych dopiero wtedy, gdy temperatura sięga 10 000°C.

Do doświadczalnego badania wpływu czynników jonizujących może służyć obwód elektryczny, zawierający kondensator powietrzny *K*, baterię ogniw *B* i galwanoskop *G* (rys. 20.18 i 20.19). W normalnych warunkach w takim obwodzie prąd stały nie popłynie ze względu na przerwę powietrzną między okładkami kondensatora. Wystarczy jednak przestrzeń między okładkami kondensatora *K* naświetlić promieniami Röntgena (rys. 20.18), promieniami biegnącymi od lampy kwarcowej (a więc bogatymi w nadfiolet) lub ustawić poniżej kondensatora *K* płomień gazowy (rys. 20.19), aby galwanometr *G* wykazał przepływ prądu świadczący o jonizacji powietrza. Z wartości prądu można wnioskować o skuteczności działania poszczególnych czynników jonizujących.



Rys. 20.18



Rys. 20.19

Liczba jonów i elektronów powstających w jednostce czasu pod działaniem nawet silnych czynników jonizujących jest stosunkowo niewielka. Na przykład 1 mg soli radowej powoduje powstanie w 1 sekundzie średnio 1 jonu na 10^6 obojętnych cząsteczek powietrza w najbliższym swym otoczeniu. Ponieważ powietrze jest stale poddane działaniu promieniowania kosmicznego i ciał radioaktywnych, więc w każdej sekundzie w każdym centymetrze sześciennym powietrza, zawierającym w warunkach normalnych około 10^{19} cząsteczek, wytwarza się od 4–8 par jonów. Jony, podobnie jak obojętne cząsteczki, są w ciągłym termicznym, chaotycznym ruchu. W wyniku spotkań jonów dodatnich i ujemnych tworzą się ponownie obojętne cząsteczki. Mamy więc znów do

czynienia z procesem rekombinacji. W wyniku równoczesnego przebiegania procesu jonizacji i rekombinacji ustala się pewien stan równowagi statystycznej, charakteryzujący się tym, że ile jonów w jednostce czasu powstaje, tyle w tym samym czasie ulega rekombinacji. Szacuje się, że w warstwie powietrza bliskiej ziemi na 1 cm^3 przypada średnio około 500–900 jonów.

Ruch jonów gazowych charakteryzujemy podobnie jak w cieczach za pomocą pojęcia ruchliwości. Wobec małej lepkości gazów ruchliwość jonów gazowych jest znacznie większa od ruchliwości jonów w elektrolitach. Jest ona rzędu $(1\text{–}2)\cdot 10^{-4}\,\dfrac{\mathrm{m/s}}{\mathrm{V/m}}$. Przy stosowaniu silnych pól elektrycznych jony gazowe rozpędzają się do takich prędkości, że uderzając np. w metalowe elektrody zdolne są wybić z nich elektrony. Los tych wybitych elektronów jest inny przy anodzie, a inny przy katodzie. Anoda przyciąga je z powrotem, natomiast katoda odpycha zmuszając je do brania udziału w przewodzeniu prądu. Ruch elektronów wobec małej ich masy jest znacznie szybszy od ruchu jonów. Po przebyciu różnicy potencjałów równej U elektron uzyskuje energię kinetyczną równą

$$\tfrac{1}{2}mv^2 = eU.$$

Stąd po podstawieniu danych liczbowych na masę i ładunek elektronu wynika zależność prędkości elektronu w m/s od przebytej różnicy potencjałów U w woltach:

$$v \approx 6\cdot 10^5 \sqrt{U}.$$

Korzystając z tego wzoru można obliczać prędkości, jakie powinny mieć elektrony, aby zdolne były przy zderzeniu z obojętną cząsteczką wywołać jej jonizację. Na przykład w przypadku tlenu lub wodoru, których potencjał jonizacji wynosi 13,5 V, odpowiednia prędkość elektronu wynosi

$$v = 600\sqrt{13{,}5} \approx 2000\ \mathrm{km/s},$$

czyli jest około 1000 razy większa od prędkości pocisków artyleryjskich dział dalekiego zasięgu.

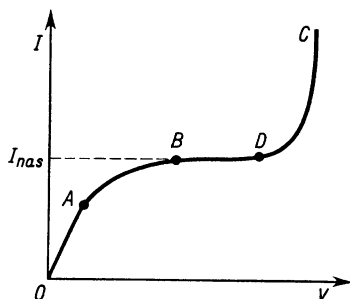
20.9.2. *Prądy w gazach pod normalnym ciśnieniem.* Przepływ prądu elektrycznego przez gazy zależy od wielu czynników, jak stopień zjonizowania, ciśnienie, temperatura, napięcie przyłożone do elektrod itp. Analizę zjawiska ograniczymy do omówienia charakteru zależności prądu (wyładowań) w gazie od napięcia przy ciśnieniu zbliżonym do normalnego (inne czynniki niezmienne) oraz do omówienia prądu (wyładowań) w gazach rozrzedzonych (p. 20.9.3).

Wykres na rys. 20.20 przedstawia zależność prądu płynącego przez gaz pod normalnym ciśnieniem od napięcia. Bierzemy pod uwagę prąd płynący w obwodzie przedstawionym na rys. 20.19. Utrzymujemy stały czynnik jonizujący, zmieniamy stopniowo napięcie między okładkami kondensatora i w ten sposób wywołujemy zmiany prądu. Początkowo (odcinek OA) przy małych napięciach mamy prostoliniowy przebieg zależności, co wskazuje na proporcjonalność prądu do przyłożonego napięcia. Innymi słowy, w tych warunkach obowiązuje prawo Ohma. Jony gazowe wytwarzane przez zewnętrzny czynnik jonizujący częściowo ulegają rekombinacji, a częściowo dochodzą

do elektrod, decydując o przepływie prądu. Ich prędkość jest proporcjonalna do natężenia pola panującego między okładkami kondensatora. Współczynnik proporcjonalności jest ruchliwością danego jonu. Na przykład

$$v^+ = u^+E.$$

Przejdźmy do analizy dalszej części wykresu na rys. 20.20. W miarę wzrostu napięcia prędkość jonów rośnie, rośnie też natężenie prądu, ale już zależność napięcia od natężenia nie jest linią prostą. Mamy odcinek krzywoliniowy *AB*. Wreszcie osiągamy takie napięcie, przy którym wszystkie jony wytwarzane np. w jednostce czasu przez czynnik jonizujący są w tym samym czasie ściągane do elektrod. Innymi słowy, w tych warunkach można zaniedbać istnienie rekombinacji. Dalsze powiększanie napięcia nie powoduje już wzrostu natężenia prądu. Mówimy, że jest osiągnięty ***prąd nasycenia*** (odcinek *BD* na wykresie). Łatwa do zmierzenia wartość natężenia prądu nasycenia jest równocześnie miarą stopnia jonizacji przestrzeni między płytami kondensatora — jest ona zależna od stosowanego czynnika jonizującego.



Rys. 20.20

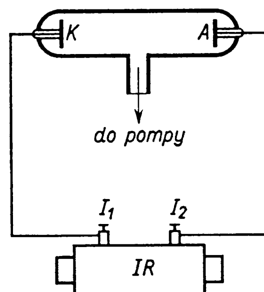
Usunięcie czynnika jonizującego w jakimkolwiek punkcie badanego przebiegu *OD* powoduje zanik przepływu prądu na skutek rekombinacji w wyniku bezładnego ruchu cieplnego. Innymi słowy, działanie czynnika jonizującego jest konieczne do podtrzymywania przepływu prądu, jeśli napięcie między okładkami kondensatora jest niewielkie. W tych warunkach mówimy o ***wyładowaniu niesamoistnym w gazie***. Jest to równocześnie wyładowanie ciemne, gdyż przepływowi prądu nie towarzyszą żadne efekty świetlne.

Jeżeli po osiągnięciu prądu nasycenia w dalszym ciągu zwiększamy napięcie, to na razie jedynym skutkiem jest wzrost prędkości ruchu jonów. Wreszcie jednak osiągnięta zostaje prędkość wystarczająca do wybicia elektronu przy zderzeniu jonu z obojętną cząsteczką gazu (*jonizacja zderzeniowa*). Od tej chwili natężenie prądu rośnie bardzo gwałtownie ze wzrostem *V* (odcinek *DC* na wykresie). W tych warunkach każdy nowo powstały jon w krótkim czasie uzyskuje w silnym polu prędkość wystarczającą do rozbicia napotkanej cząsteczki — stąd lawinowy wzrost liczby jonów, niezależny od zewnętrznego czynnika jonizującego. Usunięcie zewnętrznego czynnika jonizującego nie powoduje już zaniku przepływu prądu. Mówimy wtedy o ***wyładowaniu samoistnym*** wywołanym przez jonizację lawinową. Takiemu wyładowaniu elektrycznemu w gazie to-

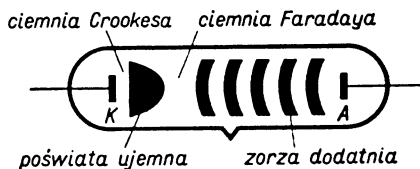
warzyszą efekty świetlne — często oślepiające światło, efekty akustyczne — trzask, efekty termiczne — wysoka temperatura.

W zależności od rodzaju źródła napięcia wywołującego pole elektryczne w przestrzeni między okładkami kondensatora, a w szczególności od jego zdolności dostarczania energii, wyładowanie samoistne może mieć charakter *iskry elektrycznej* lub *łuku elektrycznego*. Jeżeli źródło zdolne jest tylko do dostarczania pojedynczego impulsu energetycznego (powstającego np. przy nagłym rozładowaniu kondensatorów) powstaje krótkotrwałe zjawisko iskry elektrycznej o czasie trwania rzędu dziesiątków mikrosekund. Jeżeli natomiast źródło jest zdolne do trwałego dostarczania energii (np. prądnica prądu stałego) powstaje długotrwałe zjawisko łuku elektrycznego. Iskra elektryczna powstaje więc np. przy wyładowaniu piorunowym (rozładowanie kondensatora chmura-chmura lub chmura-ziemia), a łuk elektryczny np. podczas elektrycznego spawania metali.

20.9.3. *Wyładowania jarzeniowe w gazach rozrzedzonych.* Wyładowania w gazach rozrzedzonych przebiegają inaczej niż pod ciśnieniem zbliżonym do normalnego. Wyobraźmy sobie układ doświadczalny taki jak na rys. 20.21. W rurze szklanej o długości kilkudziesięciu centymetrów, połączonej z pompą próżniową, znajdują się dwie elektrody platynowe K i A. Łączymy je z iskiernikami I_1 i I_2 cewki indukcyjnej Ruhmkorffa I. R., włączając w ten sposób między elektrody A i K wysokie napięcie np. około 10 000 V. Początkowo, gdy powietrze w rurze nie jest rozrzedzone, a odległość między iskiernikami cewki mała, wyładowanie elektryczne w postaci iskry przebiega między I_1 i I_2. Gdy jednak ciśnienie w rurze spadnie do około 40 mm Hg, pojawia się w rurze świecenie w postaci ruchliwej iskry rozszerzającej się stopniowo w świecącą wstęgę przy dalszym obniżaniu ciśnienia (około 10 mm Hg). Barwa tej wstęgi zależna jest od rodzaju gazu wypełniającego rurę. W powietrzu jest ona różowofioletowa, w neonie — jaskrawo czerwona, w helu — różowa, w argonie z domieszkami par rtęci — niebieska.



Rys. 20.21



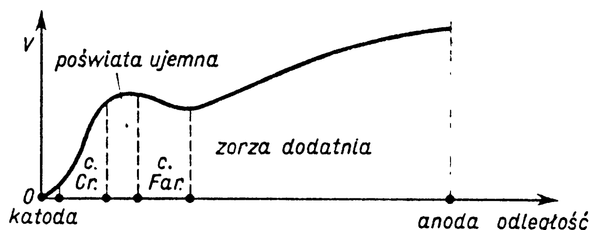
Rys. 20.22

Gdy ciśnienie w rurze spadnie do około 1–2 mm Hg, smuga świecąca urywa się od strony katody. Możemy w niej wyróżnić: zorzę dodatnią od strony anody, dalej ciemnię Faradaya i wreszcie poświatę ujemną.

Gdy ciśnienie spada poniżej 1 mm Hg w zorzy dodatniej pojawiają się uwarstwienia (warstwy jasne i ciemne) (rys. 20.22). Na katodzie powstaje świecąca warstwa katodo-

wa oddzielona ciemnią Crookesa od poświaty ujemnej. Jest rzeczą interesującą, że o długości poszczególnych warstw decyduje wysokość ciśnienia panującego w rurze. W miarę obniżania ciśnienia wszystkie warstwy się rozszerzają kosztem zorzy dodatniej, która zanika przy ciśnieniu 0,1 mm Hg. Gdy ciśnienie obniżymy do 0,001 mm Hg, świecenie wewnątrz rury w ogóle zaniknie, pojawi się natomiast fluorescencja szkła naprzeciwko katody. Będzie to już jednak skutek działania na szkło promieni katodowych.

20.9.4. *Mechanizm wyładowań jarzeniowych w gazach rozrzedzonych.* Postaramy się obecnie wniknąć nieco w istotę przebiegu wyładowań jarzeniowych. Zakładamy, że już przed przyłożeniem napięcia do elektrod rury próżniowej znajdują się w niej nieliczne jony wytworzone przez promienie kosmiczne lub ciała promieniotwórcze. Po przyłożeniu napięcia jony dodatnie dążą do katody, rozpędzają się pod wpływem pola i w przypadku dużego rozrzedzenia po przebyciu długiej drogi swobodnej z dużą energią uderzają o powierzchnię katody. Energia ta może być wystarczająca do wybicia elektronu z katody. Taki „wybity" elektron odlatuje od katody, przelatuje ruchem przyspieszonym odcinek ciemni Crookesa i zderza się z cząsteczkami gazowymi. Długość jego drogi swobodnej jest tym większa, im mniejsze jest ciśnienie gazu w rurze. W wyniku zderzeń elektronów z cząsteczkami gazu następuje wzbudzenie (a co za tym idzie i świecenie tych ostatnich — por. § 20.3), a nawet jonizacja. Te skutki działania elektronów przypadają na obszar poświaty ujemnej. W tym obszarze wytwarzają się stale nowe jony dodatnie. Mimo że część tych jonów przesuwa się w kierunku katody i powoduje wybijanie z niej dalszych elektronów — między katodą a poświatą, czyli w obszarze ciemni Crookesa, mamy najsilniejszy spad potencjału, najsilniejsze pole elektryczne.



Rys. 20.23

Ogólnie biorąc, zderzenia elektronów z cząsteczkami gazu decydują o nierównomiernym przestrzennym rozkładzie ładunku w rurze. To oczywiście ma wpływ na niejednorodność pola między elektrodami. Wykres (rys. 20.23) przedstawia rozkład potencjałów między anodą i katodą. Zaznacza się na wykresie łagodny spad potencjału w obszarze zorzy dodatniej, prawie stały potencjał na przestrzeni poświaty ujemnej i gwałtownie zaznaczony spadek w ciemni Crookesa, tworzący tzw. katodowy spad potencjału.

Wyładowania jarzeniowe znalazły szerokie zastosowanie w lampach reklamowych. Ciśnienie w nich jest tak dobrane, że prawie całą rurę wypełnia zorza dodatnia. Potrzebne napięcie dostarczane jest przez transformator prądu zmiennego.

20.10. Klasyczna teoria przewodnictwa elektrycznego metali

Poglądy badaczy na istotę przepływu prądu elektrycznego w metalach ulegały z biegiem czasu zmianie. Na przełomie wieku XIX i XX Drude opracował teorię zwaną dziś *klasyczną teorią elektronową*. Według tej teorii nośnikami prądu w metalu są elektrony walencyjne, które oderwały się od obojętnych atomów. Noszą one nazwę elektronów swobodnych i tworzą tzw. *gaz elektronowy*. Do tego „gazu elektronowego", zawartego w krystalicznej jonowej sieci przestrzennej metalu jakby w zamkniętym zbiorniku, stosowane są prawa, wynikające z kinetycznej teorii gazów, a poszczególne elektrony podlegają prawom elektrodynamiki. W każdej jednostce objętości metalu sumaryczny ładunek jonów i elektronów równa się zeru. Przy uwzględnieniu większych odległości w metalu (ujęcie makroskopowe) można więc uważać, że średnie pole elektryczne wewnętrzne też równa się zeru. Przy ujęciu mikroskopowym trzeba brać pod uwagę pola elektryczne lokalne w otoczeniu poszczególnych jonów. Ruch cieplny elektronów w pobliżu jądra atomowego ulega zakłóceniu: efekt jest do pewnego stopnia analogiczny do zderzeń występujących podczas ruchu cząsteczek gazu. Uwzględniając taką analogię można (w pierwszym przybliżeniu) mówić o zderzeniach elektronów z jonami, posługiwać się pojęciem średniej drogi swobodnej λ, średniej prędkości ruchu cieplnego u, średniej energii kinetycznej itp.

Stosunkowo proste rozważanie doprowadza do prawa Ohma. Jeśli do metalu nie jest przyłożone zewnętrzne napięcie U, ruch elektronów jest idealnie chaotyczny co do kierunku — a zatem nie ma przepływu prądu. Jeśli jednak do długości l przewodnika o przekroju S przyłożymy zewnętrzne napięcie U, to pod działaniem zewnętrznego pola elektrycznego wystąpi pewne uporządkowanie ruchu postępowego elektronów w kierunku od potencjału niższego do wyższego. Ruch ten byłby jednostajnie przyspieszony pod działaniem siły eE z przyspieszeniem $a = eE/m$, gdyby nie czynniki hamujące (np. zderzenia z jonami metalu). Zakładając, że charakter tych sił jest analogiczny do charakteru sił lepkości Stokesa (por. p. 9.8.1), czyli przyjmując proporcjonalność siły hamującej do prędkości, znajdziemy warunek przepływu prądu stałego w postaci

$$eE = \beta\bar{v},$$

gdzie $\bar{v}$ oznacza średnią prędkość ruchu elektronu, β — współczynnik proporcjonalności.

Stąd
$$\bar{v} = \frac{e}{\beta}E \quad \text{lub} \quad \bar{v} = \mu_e E. \tag{20.25}$$

Współczynnik $\mu_e = e/\beta$ jest ruchliwością elektronu i wyraża liczbowo prędkość średnią, jaką elektron uzyskuje w jednostkowym polu elektrycznym.

Natężenie prądu płynącego w przewodniku metalowym o przekroju poprzecznym S można wyrazić jako

$$I = nS\bar{v}e,$$

gdzie n oznacza liczbę elektronów w jednostce objętości. Uwzględniając równanie (20.25) i zależność (19.29a) otrzymujemy

$$I = nSe\mu_e\frac{U}{l}. \tag{20.26}$$

Z równania (20.26) widoczna jest proporcjonalność I do U (prawo Ohma). Z tego samego wzoru wynika, że opór elektryczny

$$R = \frac{l}{n\mu_e eS}.$$

Jest on zatem proporcjonalny do długości l, a odwrotnie proporcjonalny do przekroju poprzecznego S przewodnika. Przewodność właściwa metalu γ równa się

$$\gamma = \frac{1}{\varrho} = ne\mu_e.$$

Bez dowodu podamy, że z zastosowania zasady ekwipartycji energii do gazu elektronowego pozostającego w równowadze termicznej wynika dalsza zależność:

$$\gamma = \frac{e^2 n\lambda u}{6kT}, \tag{20.27}$$

gdzie u oznacza średnią prędkość ruchu termicznego, k — stałą Boltzmanna równą R/N_A (por. § 16.6). Z równania (20.27) widać, że z teorii Drudego wynika odwrotna proporcjonalność przewodności elektrycznej metali do temperatury T.

Dalsze rozwinięcie teorii Drudego prowadzi do *prawa Wiedemanna–Franza*, ustalonego pierwotnie na podstawie doświadczeń, według którego *stosunek przewodności cieplnej właściwej do przewodności elektrycznej właściwej metali jest stały w danej temperaturze*:

$$\frac{\text{przewodność cieplna właściwa}}{\text{przewodność elektryczna właściwa}} = \alpha T,$$

gdzie α oznacza tzw. stałą Wiedemanna–Franza. Eksperymentalnie znaleziona wartość liczbowa tego współczynnika dla większości metali w temperaturze pokojowej wynosi około $24 \cdot 10^3 \frac{\mathrm{J \cdot \Omega}}{\mathrm{m \cdot s \cdot K^2}}$. Z teorii Drudego wynika, że wartość tego współczynnika jest równa

$$\alpha = \frac{3k^2}{e^2}$$

(znaczenie symboli takie, jak poprzednio). Wynik eksperymentalny uzyskany w temperaturze pokojowej i wynik teoretyczny wykazują niezłą zgodność.

Dokładniejsze badania wykazały, że prawo Wiedemanna–Franza obowiązuje z możliwym przybliżeniem tylko w temperaturach pokojowych i wyższych. W temperaturach niskich α traci swą „stałość" dla różnych metali. Tych zjawisk, jak również szeregu innych, teoria klasyczna nie była w stanie wyjaśnić, a w niektórych przypadkach prowadziła do wyników sprzecznych z doświadczeniem.

Dalszy rozwój poglądów na sprawę przewodnictwa elektrycznego metali przypada na lata dwudzieste bieżącego stulecia i wiąże się z nazwiskami takich badaczy jak Sommerfeld, Frenkiel i inni. Podstawowe wiadomości z kwantowej teorii przewodnictwa omówimy w § 32.1–32.8.

Pytania i zadania

1. Jaka jest definicja natężenia prądu elektrycznego i jego jednostki w układzie SI?

2. Czy mechanizm przepływu prądu w metalach, elektrolitach i gazach jest jednakowy?

3. Jaką wielkość fizyczną definiuje się na podstawie prawa Ohma? Podać definicję tej wielkości i jej jednostki.

4. Jaka jest zależność oporu elektrycznego przewodnika od jego rozmiarów geometrycznych? Jaka wielkość charakteryzuje zależność oporu przewodnika od rodzaju materiału?

5. Jakie czynniki wpływają na opór przewodników metalicznych?

6. Jak są ze sobą związane: a) opór przewodnika i jego przewodność, b) opór właściwy przewodnika i jego przewodność właściwa? Podać nazwy jednostek przewodności i przewodności właściwej w układzie SI.

7. Co rozumiemy przez gęstość prądu elektrycznego? Jaki jest związek między gęstością prądu i natężeniem pola elektrycznego?

8. Podać treść praw Kirchhoffa i przeanalizować je na przykładzie węzłów i oczek wydzielonych w obwodzie mostka Wheatstone'a.

9. Na czym polega łączenie szeregowe i równoległe oporów i do jakich wyników prowadzi każde z nich? Podać przykłady zastosowania każdego rodzaju łączenia oporów.

10. Omówić szeregowe i równoległe łączenie źródeł prądu i wpływ obu sposobów łączenia na wypadkową siłę elektromotoryczną.

11. Jakie wzory wyrażają pracę (energię) i moc prądu elektrycznego?

12. Jakie zjawisko nazywamy elektrolizą? Jakie prawa nim rządzą?

13. Podać definicje równoważnika elektrochemicznego i stałej Faradaya.

14. Jaki jest mechanizm elektrolizy?

15. Jaką stałą fizyczną można wyznaczyć znając stałą Faradaya i stałą Avogadra? Jaką inną metodą ta sama wielkość była wyznaczana? Z jakim prawem dotyczącym ruchu ciał w ośrodkach lepkich wiąże się ta metoda?

16. Jakie są podstawowe zastosowania elektrolizy?

17. Jaki jest mechanizm przepływu prądu w gazach? Jakie są podstawowe czynniki jonizujące? Co to jest energia i potencjał jonizacji?

18. W jakich warunkach prawo Ohma stosuje się do elektrolitów i gazów?

19. Jak się przedstawia przepływ prądu elektrycznego przez gaz przy stopniowym zmniejszaniu ciśnienia w doświadczeniu przedstawionym na rys. 20.21?

20. Jakie są założenia klasycznej teorii przewodnictwa metali? W jaki sposób w oparciu o tę teorię wyprowadza się prawo Ohma? Co jest treścią prawa Wiedemanna-Franza?

21. Dwa przewodniki o jednakowej średnicy, jeden z węgla, a drugi z miedzi, są połączone szeregowo. Znając opory właściwe i współczynniki temperaturowe oporu obu substancji (por. tab. 20.1) wyjaśnić dlaczego opór wypadkowy może nie zależeć od temperatury i obliczyć, przy jakim stosunku długości obu przewodników to nastąpi.

Odp. $\frac{L_m}{L_w} = \frac{\varrho_w \alpha_w}{\varrho_m \alpha_m}$.

22. Opór wypadkowy dwóch przewodników połączonych szeregowo wynosi a Ω, a połączonych równolegle b Ω. Obliczyć opory tych przewodników.

Odp. $\frac{a+\sqrt{a(a-4b)}}{2}\,\Omega,\ \frac{a-\sqrt{a(a-4b)}}{2}\,\Omega$.

23. Obliczyć prąd I płynący w obwodzie przedstawionym na rys. 20.24, jeżeli $R_1 = 9\ \Omega$, $R_2 = R_3 = 5\ \Omega$, $R_4 = R_5 = 20\ \Omega$, siła elektromotoryczna źródła $E = 24$ V. Opór wewnętrzny źródła pominąć.

Odp. 1 A.

24. Jaki opór należy połączyć szeregowo z baterią złożoną z dziesięciu szeregowo połączonych aku-

mulatorów, o jednakowych SEM-nych E_1 = 1,9 V i oporze wewnętrznym R_w = 0,01 Ω, aby prąd ładowania płynący pod stałym napięciem E_2 = 220 V równał się 4 A.

Odp. 50,15 Ω

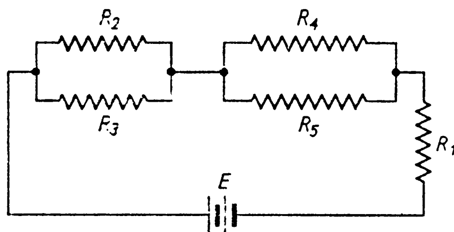
25. W jaki sposób można n-krotnie rozszerzyć zakres skali woltomierza o oporze wewnętrznym R?

26. W celu wyznaczenia oporu R utworzono dwa obwody. W obwodzie jak na rys. 20.25a amperomierz wskazuje I_1 = 2,15 A, woltomierz zaś U_1 = 49,5 V; w obwodzie jak na rys. 20.25b amperomierz wskazuje I_2 = 2 A, a woltomierz U_2 = 50 V. Ile wynosi opór R?

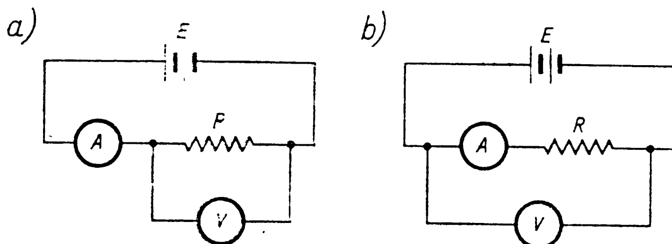
Odp. 24,77 Ω.

27. Froterkę o mocy P = 200 W na napięcie U_1 = 110 V chcemy włączyć do sieci o napięciu U = = 220 V. Jak i jaki opór należy włączyć do obwodu froterki?

Odp. szeregowo 60,5 Ω.



Rys. 20.24



Rys. 20.25

28. Ile czasu wymaga zagotowanie 1 kg wody o temperaturze początkowej t = 20°C przy użyciu grzejnika elektrycznego, którego spirala ma opór R = 24 Ω, jeżeli sprawność grzejnika wynosi 80%, a napięcie sieci 220 V?

Odp. 3,4 min.

29. Do miedzianego drutu o długości L = 40 m i średnicy $2r$ = 1 mm przyłożono napięcie U = = 1 V. Obliczyć: natężenie płynącego prądu, jego gęstość, natężenie pola elektrycznego oraz szybkość wydzielania ciepła.

30. Poniklowanie przedmiotu o powierzchni S = 100 cm^2 trwało t = 10 h przy przepływie prądu I = 0,5 A. Jaka jest grubość warstwy niklu? Masa atomowa niklu wynosi 58,7, jego wartościowość w = = 2, a gęstość ϱ = 8,8 g/cm^3.

Odp. 0,006 cm.

ROZDZIAŁ 21

Pole magnetyczne

21.1. Wiadomości wstępne

Oddziaływania magnetyczne odkryto wcześniej niż oddziaływania elektryczne. Wiąże się to z istnieniem w przyrodzie tzw. magnesów trwałych (np. ruda żelazna — magnetyt), jak również z tym, że Ziemia zachowuje się jak wielki magnes. Magnesy trwałe wywierają działanie na żelazo, stal (sztuczne magnesy stalowe); znalazły one zastosowanie jako wskaźniki kierunku północnego i południowego na Ziemi (kompasy).

Z historią rozwoju magnetyzmu, a w latach późniejszych elektromagnetyzmu, wiążą się m. in. nazwiska Coulomba (1785 r. — prawo wzajemnego oddziaływania biegunów magnetycznych), Oersteda, Ampère'a, Biota i Savarta (pole magnetyczne prądu elektrycznego), Faradaya i Lenza (indukcja elektromagnetyczna). Wybierając przykładowo parę nazwisk z ostatnich dziesiątków lat można wymienić Lawrence'a (cyklotron), Powella i Elsassera (pole magnetyczne ziemskie), Blocha i Purcella (magnetyczny rezonans jądrowy).

W okresie początkowym rozwoju magnetyzmu wprowadzono pojęcie mas magnetycznych: północnej i południowej (lub dodatniej i ujemnej), stwierdzając równocześnie niemożliwość ich rozdzielenia (zasadnicza różnica w stosunku do ładunków elektrycznych dodatnich i ujemnych). Z biegiem czasu pojęcie masy magnetycznej wyszło z użycia; obecnie nie traktujemy już sił magnetycznych jako skutku istnienia mas magnetycznych. Posługujemy się jednak w dalszym ciągu pojęciem biegunów magnetycznych, mając na myśli takie obszary w pobliżu końców magnesów trwałych (w postaci sztabek, podków itp.) lub elektromagnesów, w których dają się zauważyć najsilniejsze oddziaływania magnetyczne (np. jeśli magnes sztabkowy zbliżymy do opiłków żelaznych, to bieguny magnetvczne przyciągają ich najwięcej).

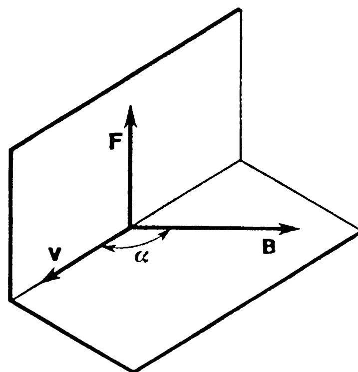
Istnienie pól magnetycznych jest traktowane obecnie jako objaw wtórny, jako skutek ruchu ładunków elektrycznych. Wszelki przepływ prądu elektrycznego powoduje powstanie pola magnetycznego. Jest to niezależne od natury prądu: może to być prąd elektronowy w przewodniku metalicznym, prąd jonowy w elektrolicie, prąd w gazie. Pole magnetyczne towarzyszy też ruchowi elektronów w atomie, ruchowi jąder atomowych w cząsteczkach itd.

Do charakterystyki wektorowej pola magnetycznego wykorzystuje się, podobnie jak w elektrostatyce, trzy podstawowe wektory, a mianowicie: wektor indukcji magnetycznej **B**, wektor natężenia pola magnetycznego **H** oraz wektor namagnesowania **M**. Dla silniejszego zaakcentowania bliskiej więzi elektryczności z magnetyzmem na plan pierwszy wysuniemy wektor **B**.

21.2. Indukcja magnetyczna

Indukcję magnetyczną **B** definiujemy wykorzystując siłę oddziaływania pola magnetycznego na poruszający się w tym polu dodatni ładunek próbny Q_0. Doświadczenie uczy, że siła **F** działająca w danym punkcie pola na ładunek Q_0 poruszający się z prędkością **v**: 1) zmienia swą wartość od 0 do F_{max} w zależności od kierunku **v**, 2) działa zawsze w kierunku prostopadłym do **v**, tzn. nie usiłuje przyspieszyć lub zahamować ładunku Q_0 na jego pierwotnym torze (stycznym do **v**), lecz usiłuje przesunąć go w bok od pierwotnego kierunku ruchu. Stwierdzenie w przypadku poruszającego się ładunku istnienia siły o wymienionych cechach świadczy o tym, że ruch ładunku Q_0 odbywa się w polu magnetycznym. Wektorowi indukcji magnetycznej **B** tego pola przypisujemy — według umowy — kierunek zgodny z tym kierunkiem prędkości **v**, przy którym siła $\mathbf{F} = 0$. Pełną definicję wektora **B** przedstawia równanie

$$\mathbf{F} = Q_0(\mathbf{v}\times\mathbf{B}) \quad \text{lub} \quad F = Q_0 vB\sin(\mathbf{v},\mathbf{B}). \tag{21.1}$$



Rys. 21.1

Z właściwości iloczynu wektorowego (por. § 6 Wiadomości wstępnych) wynika, że:

1. Trzy wektory **F**, **v**, **B** stanowią taki układ, że siła **F** jest prostopadła do płaszczyzny wektorów **v** i **B** (rys. 21.1).

2. Przy zmianie kierunku **v**, bez zmiany jej wartości liczbowej, siła zmienia swą wartość od 0 (gdy $\mathbf{v}\parallel\mathbf{B}$) do F_{max} (gdy $\mathbf{v}\perp\mathbf{B}$).

3. Maksymalna wartość siły F wynosi

$$F_{max} = Q_0 vB \quad (\text{gdy } \mathbf{v}\perp\mathbf{B}). \tag{21.2}$$

Z równania (21.2) wynika, że

$$B = \frac{F_{max}}{Q_0 v}. \tag{21.3}$$

To równanie określa wartość liczbową indukcji magnetycznej **B**.

Z równania (21.3) znajdujemy, *że jednostką* **B** *jest* $\frac{N}{C \cdot m/s}$ lub $\frac{N}{A \cdot m}$. Jednostce tej nadano nazwę *tesla* (skrót T), czyli

$$1\,T = 1\frac{N}{A \cdot m}.$$

Z wektorem indukcji magnetycznej kojarzymy:

a) pojęcie *linii indukcji magnetycznej* (linia indukcji jest w każdym swym punkcie styczna do kierunku **B**),

b) *gęstość linii indukcji magnetycznej* (przez jednostkę powierzchni prostopadłą do **B** przechodzi tyle linii, ile wynosi wartość liczbowa **B** na tej powierzchni),

c) *strumień indukcji* Φ (określony jako $\int B dS$). Jednostką strumienia Φ jest $T \cdot m^2$. Jednostce tej nadano nazwę *weber* (skrót Wb).

$$1\,Wb = 1\,T \cdot m^2.$$

Wszystkie wymienione wyżej pojęcia i wielkości mają swoje odpowiedniki elektrostatyczne odniesione do wektorów natężenia pola elektrycznego **E** i indukcji elektrostatycznej **D**.

Jeśli ładunek próbny Q_0 porusza się w obszarze, w którym istnieją równocześnie pole elektryczne o natężeniu **E** i pole magnetyczne o indukcji **B**, to wypadkowe oddziaływanie obu pól sprowadza się do siły **F**:

$$\mathbf{F} = Q_0\mathbf{E} + Q_0(\mathbf{v} \times \mathbf{B}). \tag{21.4}$$

Siłę tę nazywamy często *siłą Lorentza*.

Pytanie. 1. Przy przelocie pewnych naładowanych cząstek przez pole magnetyczne stwierdzono, że siła odchylająca te cząstki od pierwotnego kierunku prędkości wynosi

a) $\mathbf{F} = e(\mathbf{v} \times \mathbf{B})$,

b) $\mathbf{F} = -e(\mathbf{v} \times \mathbf{B})$.

Jakie to mogły być cząstki? Czy siły działające w obu przypadkach były identyczne, jeśli nie — sprecyzować, na czym polegała różnica.

2. Czy wyrażenia $\mathbf{F} = e(\mathbf{v} \times \mathbf{B})$ i $\mathbf{F} = -e(-\mathbf{v} \times \mathbf{B})$ przedstawiają siły jednakowe, czy też różne? Opisać dokładnie rodzaje możliwych cząstek poruszających się i wskazać różnice w ich ruchach.

3. Podać postać wzoru na siłę Lorentza a) dla protonu, b) dla elektronu.

21.3. Działanie pola magnetycznego na odcinek przewodnika z prądem elektrycznym

Wiemy już, że pole magnetyczne o indukcji **B** działa na pojedynczy ładunek Q_0, poruszający się w tym polu siłą **F** wyrażoną wzorem

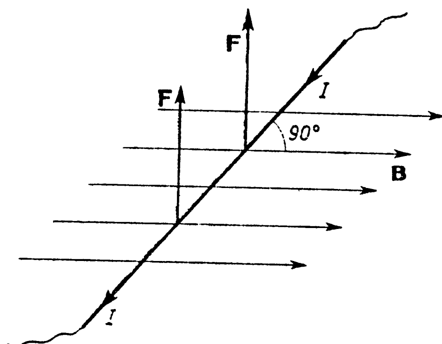
$$\mathbf{F} = Q_0(\mathbf{v} \times \mathbf{B}).$$

Obecnie zajmiemy się badaniem siły, jaką pole magnetyczne działa na odcinek przewodnika, przez który płynie prąd elektryczny. Przede wszystkim stwierdzimy, że na podstawie takich badań nie można ustalić charakteru nośników prądu.

Pamiętamy, że w przewodniku metalowym nośnikami prądu są swobodne elektrony o ładunku $-e$. Poruszają się one od potencjału niższego do wyższego, a więc w kierunku przeciwnym względem kierunku przyjmowanego normalnie za kierunek przepływu prądu. Łatwo sprawdzić, że podstawienie do wzoru na siłę wielkości związanych z ruchem rzeczywistych nośników (czyli elektronów o ładunku $-e$ i prędkości $-\mathbf{v}$) da wynik identyczny z tym, jaki otrzymalibyśmy odnosząc wzór do nośników o ładunku $+e$, mających prędkość przeciwnie skierowaną $(+\mathbf{v})$ (por. pytanie 2 z § 21.2):

$$\mathbf{F} = -e(-\mathbf{v}\times\mathbf{B}) = e(\mathbf{v}\times\mathbf{B}).$$

Innymi słowy, badanie siły działającej na przewodnik z prądem w polu magnetycznym nie pozwala stwierdzić charakteru nośników prądu.



Rys. 21.2

W celu obliczenia siły pochodzącej od jednorodnego pola i działającej na odcinek l przewodu, przez który płynie prąd I, rozważmy początkowo przypadek, gdy przewodnik umieszczony jest prostopadle do $\mathbf{B}$ (rys. 21.2). W tych warunkach siła $\mathbf{F}'$, działająca na każdy z nośników prądu, będzie jednakowo skierowana (prostopadle do $\mathbf{v}$ i do $\mathbf{B}$) i równa $F' = evB$. A zatem siła wypadkowa będzie równa sumie arytmetycznej sił działających na wszystkie nośniki znajdujące się w rozważanym odcinku przewodu. Przyjmując, że gęstość nośników prądu (liczba nośników w jednostce objętości) jest n, znajdziemy, że ogólna ich liczba w odcinku l przewodnika o przekroju S wynosi nlS. Siła wypadkowa jest więc równa

$$F = nlSF' = nlSe\bar{v}B.$$

Warto tu podkreślić, że występująca we wzorze prędkość $\bar{v}$ jest prędkością średnią ruchu poszczególnych nośników prądu (a nie np. prędkością ruchu przewodnika jako całości). Ale w § 20.10 stwierdziliśmy, że natężenie prądu w przewodniku metalowym spełnia zależność

$$I = nS\bar{v}e.$$

Średnia prędkość unoszenia nośników $\bar{v}$ odpowiada prędkości $\bar{v}$ z ostatniego wzoru, a zatem

$$F = IlB. \tag{21.5}$$

Wzór ten wyraża siłę w przypadku prostopadłego ustawienia l i **B**. W przypadku ogólnym prostoliniowego przewodnika o długości l tworzącego dowolny kąt z wektorem indukcji magnetycznej **B** siła wyraża się wzorem

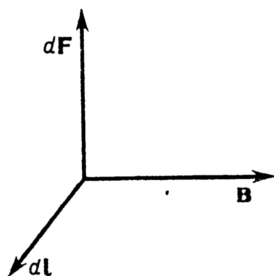
$$\mathbf{F} = I(\mathbf{l} \times \mathbf{B}). \tag{21.6}$$

Wprowadzone w tym wzorze oznaczenie **l** przedstawia wektor o wartości liczbowej równej długości prostoliniowego odcinka przewodu, o kierunku zgodnym z tym przewodem i o zwrocie wyznaczonym przez kierunek przepływu prądu, tzn. przez kierunek ruchu ładunków dodatnich. Warto w tym miejscu wrócić do rys. 21.2 i sprawdzić zaznaczony tam kierunek i zwrot siły **F**.

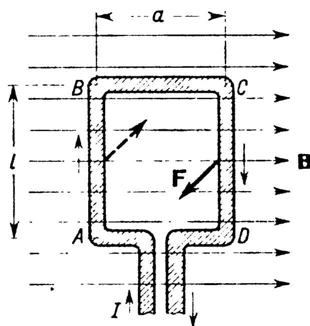
W odniesieniu do elementu przewodnika o długości dl, umieszczonego w dowolnym polu magnetycznym, wzór (21.6), przyjmuje postać

$$d\mathbf{F} = I(d\mathbf{l} \times \mathbf{B}). \tag{21.7}$$

Zestawienie w przestrzeni kierunków $d\mathbf{l}$, **B**, $d\mathbf{F}$ przedstawia rys. 21.3.



Rys. 21.3



Rys. 21.4

Całkowanie wzoru (21.7) — mniej lub bardziej skomplikowane zależnie od kształtu przewodnika — pozwala obliczyć siłę wypadkową, działającą na przewodnik krzywoliniowy.

Siłę wywołaną przez pole magnetyczne i działającą na przewodnik z prądem, określoną równaniami (21.6) i (21.7), nazywamy *siłą elektrodynamiczną*.

P r z y k ł a d. Jakie będą skutki działania jednorodnego pola magnetycznego o indukcji **B** na swobodnie osadzoną prostokątną ramkę $ABCD$ z drutu o długościach boków l i a (rys. 21.4). Przez ramkę płynie prąd I o kierunku zaznaczonym na rysunku.

Jeśli w pierwszej chwili powierzchnia ramki jest równoległa do linii indukcji **B**, to na jej boki AB i CD działają siły prostopadłe do płaszczyzny rysunku, tworzące parę sił. Moment tej pary sił, zgodnie ze wzorami (21.5) i (6.5), jest równy liczbowo

$$M_{\text{pary}} = IlBa. \tag{21.8}$$

Moment ten usiłuje obrócić ramkę. Jeśli ramka jest swobodna, to rzeczywiście wykonuje obrót, w czasie obrotu jednak moment obrotowy stopniowo maleje i staje się równy zeru przy położeniu ramki prostopadłym do linii sił pola. W takim położeniu siły działające na wszystkie cztery boki ramki dążą do jej rozciągnięcia, lecz nie nadają jej ruchu obrotowego.

P y t a n i e. 1. Wykazać, że w każdym położeniu pośrednim moment pary sił elektrodynamicznych, działających na ramkę, wyraża się wzorem

$$M_{\text{pary}} = IlBa\sin(\mathbf{n}, \mathbf{B}). \qquad (21.9a)$$

Kąt stanowiący argument sinusa we wzorze (21.9a) jest kątem między wektorem jednostkowym **n** o kierunku normalnej do powierzchni ramki i wektorem **B**.

2. Dlaczego przy położeniu ramki równoległym do linii sił pola na boki BC i AD nie działają siły elektrodynamiczne?

3. Dlaczego można zaniedbywać działanie pola na 2 przewody równoległe: doprowadzające i odprowadzające prąd z ramki przy umieszczeniu ich blisko siebie?

Wróćmy raz jeszcze do wzoru (21.9a). Występujący tam iloczyn la przedstawia pole S przekroju ramki, a zatem

$$M_{\text{pary}} = IBS\sin(\mathbf{n}, \mathbf{B}). \qquad (21.9b)$$

Ze wzoru zniknęły wielkości związane z kształtem ramki, pozostało tylko pole S jej powierzchni. Można wykazać doświadczalnie, że rzeczywiście kształt zamkniętego obwodu nie odgrywa roli w rozważanym zagadnieniu. Ramki, np. prostokątna i kołowa, o tej samej powierzchni S i opływane przez ten sam prąd I — w tym samym polu o indukcji **B** podlegają przy jednakowym ustawieniu względem **B** temu samemu momentowi obrotowemu.

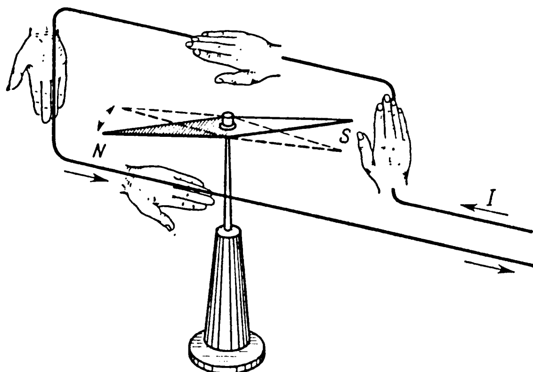
Jeśli badany obwód złożony jest z N równoległych, bliskich siebie ramek, to wypadkowy moment obrotowy jest N razy większy, gdyż momenty działające na poszczególne ramki sumują się arytmetycznie.

Zdolność wywoływania obrotu obwodów zamkniętych, umieszczonych w polu magnetycznym, została wykorzystana w licznych urządzeniach. Wymienić tu można silniki elektryczne (budowa i działanie uwzględnione w kursie elektrotechniki), jak również liczne przyrządy pomiarowe prądu i napięcia elektrycznego.

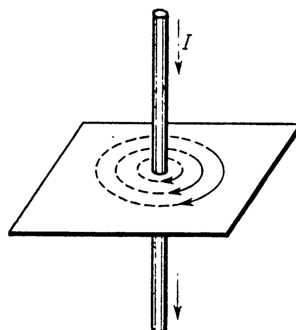
21.4. Pole magnetyczne w otoczeniu przewodu z prądem

Jak już sygnalizowaliśmy w § 21.1, przepływowi prądu elektrycznego (a więc ruchowi ładunków elektrycznych) zawsze towarzyszy pole magnetyczne. Fakt ten został po raz pierwszy stwierdzony w doświadczeniu Oersteda (1820 r.). W doświadczeniu tym istnienie pola magnetycznego w otoczeniu przewodnika liniowego z prądem zostało stwierdzone przy użyciu małej, podpartej na ostrzu igiełki magnetycznej, stanowiącej *dipol magnetyczny* o biegunach N (północny) i S (południowy). Wiemy, że igiełka magnetyczna wsparta na ostrzu ustawia się w płaszczyźnie południka magnetycznego Ziemi. Jeśli równolegle do igiełki rozciągniemy przewodnik (rys. 21.5) i przepuścimy przez niego prąd elektryczny, to igiełka zmieni swe ustawienie. Obok działania składowej poziomej natężenia pola magnetycznego ziemskiego wystąpi moment pary sił po-

chodzących od pola magnetycznego prądu i skręcających igiełkę. Kierunek skręcenia igiełki można określać w oparciu o różne reguły, np. o regułę pływaka lub regułę prawej dłoni. Według tej ostatniej, jeśli prawą dłoń umieścimy wzdłuż przewodnika, palcami w kierunku płynięcia prądu, a dłonią w stronę igiełki, to odchylony wielki palec wskaże nam kierunek skręcenia bieguna północnego igiełki.

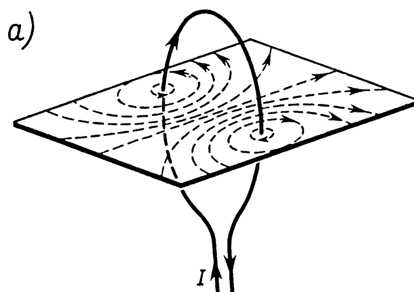


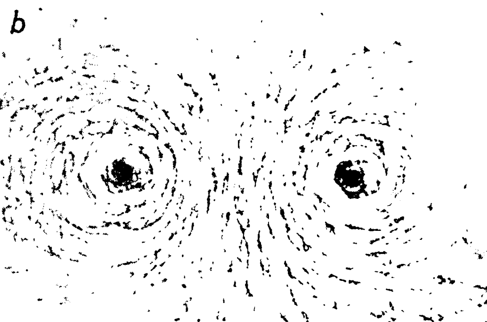
Rys. 21.5



Rys. 21.6

Rysunki 21.6 i 21.7 przypominają znane z kursu szkolnego doświadczenia, w których wystąpienie uporządkowanego rozkładu opiłków żelaznych (małe dipole magnetyczne) w płaszczyznach prostopadłych do przewodnika nie tylko potwierdza istnienie pola magnetycznego w otoczeniu przewodnika liniowego i kołowego z prądem, ale ponadto pokazuje kształt linii sił pola w obu przypadkach. Warto zapamiętać, że przebieg linii sił pola magnetycznego w otoczeniu nieskończenie długiego liniowego przewodnika z prądem jest kołowy. Linie sił leżą w płaszczyznach prostopadłych do przewodnika, środki wszystkich kół znajdują się na osi przewodnika.





Rys. 21.7

Zwrot linii sił określamy np. według *reguły korkociągu*: zwrot linii sił pola jest zgodny z kierunkiem, w jakim należy obracać rączkę korkociągu, umieszczonego wzdłuż przewodnika, aby uzyskać ruch postępowy korkociągu zgodny z kierunkiem prądu.

Rozpatrzmy od strony ilościowej przypadek pola magnetycznego w otoczeniu nie-

skończenie długiego prostoliniowego przewodu z prądem. Mierząc indukcję **B** w otoczeniu takiego przewodnika w warunkach: 1) stałej odległości r od przewodnika i zmienianej wartości I, a następnie 2) stałego prądu I i zmienianej odległości r — można stwierdzić, że indukcja B jest wprost proporcjonalna do prądu, a odwrotnie — do odległości od przewodu ($B \sim I/r$). Wobec tego, że w wymienionej zależności nie ma żadnej swobody w wyborze jednostki którejkolwiek z wielkości, współczynnik proporcjonalności, jaki musimy uwzględnić przechodząc od proporcjonalności do równości ma ściśle określoną wartość równą $\mu_0/2\pi$. Wielkość μ_0 nosi nazwę *przenikalności magnetycznej próżni* i wynosi $4\pi \cdot 10^{-7}$ Wb/A · m (lub T · m/A). Wartość liczbowa μ_0 wiąże się ściśle z definicją ampera (por. § 3 „Wiadomości wstępnych"). A zatem indukcja magnetyczna w otoczeniu nieskończenie długiego prostoliniowego przewodu z prądem wyraża się wzorem

$$B = \frac{\mu_0 I}{2\pi r}. \tag{21.10}$$

Jest to zgodne z ogólniejszym prawem, zwanym *prawem Ampère'a*, według którego dla dowolnego pola magnetycznego i dowolnej zamkniętej drogi całkowania obejmującej powierzchnię przebijaną przez całkowity prąd I

$$\oint B dl = \mu_0 I. \tag{21.11}$$

W rozważanym przypadku wykonujemy obieg po zamkniętym obwodzie kołowym, utworzonym przez dowolną linię indukcji, przy czym w każdym punkcie tego obwodu: 1) kierunki **B** i d**l** są zgodne, 2) **B** ma wartość stałą. A zatem

$$B \cdot 2\pi r = \mu_0 I,$$

co jest równoważne wyrażeniu (21.10).

Trzeba tu od razu podkreślić, że obliczenie indukcji B na podstawie prawa wyrażonego wzorem (21.11) wiąże się czasem z tak skomplikowanym całkowaniem, że szukamy innych dróg rozwiązania problemu. Pomocne okazuje się tu prawo Biota i Savarta.

21.5. Prawo Biota i Savarta

Niech CD (rys. 21.8) przedstawia odcinek długiego krzywoliniowego przewodnika, przez który płynie prąd I. Dla obliczenia indukcji magnetycznej **B** w punkcie A dzielimy przewodnik na nieskończenie małe elementy d**l**, traktując je jako wektory o zwrocie zgodnym ze zwrotem I. Jeden z takich elementów zaznaczony jest na rys. 21.8. Jego odległość od punktu A wynosi **r** (zwrot wektora **r** od elementu przewodnika do punktu A). Nieskończenie mały element dl przewodnika wytwarza w punkcie A indukcję dB, wyrażoną wzorem Biota i Savarta, a mianowicie

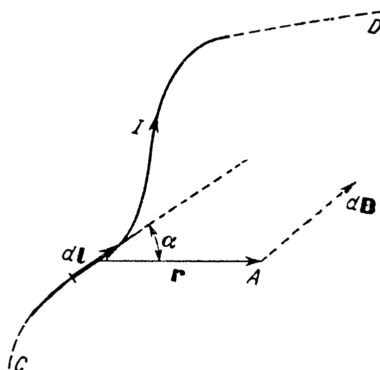
$$dB = \frac{\mu_0 I dl \sin\alpha}{4\pi r^2}, \tag{21.12}$$

gdzie α oznacza kąt między kierunkiem $d\mathbf{l}$ i $\mathbf{r}$. To wyrażenie przedstawia tylko wartość liczbową $d\mathbf{B}$. Przechodząc do postaci wektorowej napiszemy

$$d\mathbf{B} = \frac{\mu_0 I}{4\pi r^3}(d\mathbf{l}\times\mathbf{r}). \tag{21.13}$$

A zatem ujmując słownie treść wzorów (21.12) i (21.13) powiemy, że

1. Wartość liczbowa indukcji $d\mathbf{B}$ wywołanej przez element $d\mathbf{l}$ przewodnika jest proporcjonalna do natężenia prądu I, do długości elementu dl, odwrotnie proporcjonalna do kwadratu odległości r i zależna od kąta α utworzonego przez kierunki $d\mathbf{l}$ i $\mathbf{r}$.



Rys. 21.8

2. Kierunek i zwrot $d\mathbf{B}$ jest zgodny z kierunkiem i zwrotem iloczynu wektorowego $d\mathbf{l}\times\mathbf{r}$. Można też zwrot $d\mathbf{B}$ określać *regułą pływaka*: pływak płynący wzdłuż przewodnika zgodnie z kierunkiem prądu i zwrócony twarzą do rozważanego punktu wyciągniętą lewą ręką wskazuje zwrot $d\mathbf{B}$. Całkowita indukcja $\mathbf{B}$ wytworzona w punkcie A dzięki przepływowi prądu w całym przewodniku jest sumą geometryczną wektorów $d\mathbf{B}$ wytworzonych przez wszystkie elementy $d\mathbf{l}$ przewodnika (a zatem jest to całka wektorowa $\mathbf{B} = \int d\mathbf{B}$).

Jak już mówiliśmy, wyznaczenie całkowitej indukcji $\mathbf{B}$ pochodzącej od wszystkich elementów przewodnika jest na ogół trudne. W pewnych szczególnych przypadkach, a mianowicie wtedy gdy cały obwód leży w jednej płaszczyźnie, rozumowanie jest stosunkowo proste.

Rozpatrzmy dla przykładu indukcję magnetyczną 1) w środku obwodu kołowego, 2) w odległości r_0 od nieskończenie długiego prostoliniowego przewodnika.

1. Przez obwód kołowy o promieniu r płynie prąd o natężeniu I (rys. 21.9). Wszystkie elementy tego obwodu wywołują w jego środku indukcję $d\mathbf{B}$ o jednakowych kierunkach i zwrotach. A zatem sumowanie wektorowe $d\mathbf{B}$ sprowadza się do sumowania aryt-

metycznego. Uwzględniając nieskończenie wiele małych elementów dl tworzących w całości obwód kołowy znajdziemy, że

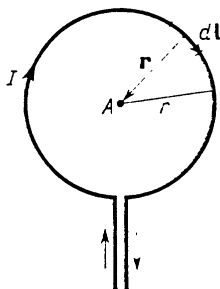
$$B = \oint dB = \oint \frac{\mu_0 I}{4\pi r^2} dl,$$

gdyż $\sin(d\mathbf{l}, \mathbf{r}) = 1$ dla każdego elementu. Stąd

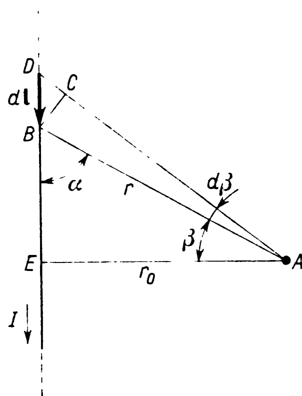
$$B = \frac{\mu_0 I}{4\pi r^2} 2\pi r = \frac{\mu_0 I}{2r}. \tag{21.14}$$

W przypadku obwodu kołowego złożonego z n gęsto nawiniętych zwojów indukcja magnetyczna w środku obwodu wynosi

$$B = \frac{\mu_0 I n}{2r}. \tag{21.15}$$



Rys. 21.9



Rys. 21.10

2. Nieco bardziej skomplikowane jest obliczenie indukcji magnetycznej w punkcie A, leżącym w odległości r_0 od nieskończenie długiego, prostoliniowego przewodnika (rys. 21.10). Tym razem indukcję B możemy przedstawić jako

$$B = \int_{-\infty}^{+\infty} \frac{\mu_0 I dl \sin\alpha}{4\pi r^2},$$

gdyż sumowanie wektorowe wszystkich indukcji $d\mathbf{B}$ pochodzących od nieskończenie małych elementów $d\mathbf{l}$ przewodnika można zastąpić zwykłym sumowaniem arytmetycznym w związku z tym, że kierunki i zwroty wszystkich wektorów $d\mathbf{B}$ są jednakowe (w przypadku przedstawionym na rys. 21.10 — prostopadłe do płaszczyzny rysunku w górę). Łączymy punkt A z końcami elementu $dl = BD$. Odległość BA oznaczamy przez r. Kąt między r_0 i r oznaczamy przez β, między BA i DA przez $d\beta$, kąt zaś EBA przez α. Postaramy się dl wyrazić za pośrednictwem r_0, r i kąta β. W tym celu z punktu B spuszczamy na DA prostopadłą BC, co jest w rozważanym przypadku prawie równo-

znaczne z zakreśleniem łuku BC promieniem r. Z ogólnej zależności między kątem, łukiem i promieniem wynika, że

$$BC = r d\beta.$$

Z trójkąta DBC

$$dl = \frac{BC}{\cos DBC},$$

ale kąt DBC równa się β (kąty o ramionach wzajemnie prostopadłych). A zatem

$$dl = \frac{r d\beta}{\cos\beta}.$$

Szukamy wyrażenia dl/r^2:

$$\frac{dl}{r^2} = \frac{d\beta}{r\cos\beta}.$$

Z trójkąta EAB wynika zależność

$$r_0 = r\cos\beta,$$

czyli

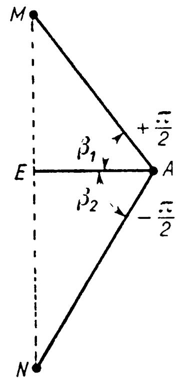
$$\frac{dl}{r^2} = \frac{d\beta}{r_0}.$$

Uwzględniając równość $\sin\alpha = \cos\beta$ i wracając do całki wyrażającej indukcję **B** w punkcie A otrzymujemy

$$B = \frac{\mu_0 I}{4\pi r_0}\int\limits_{\beta_2}^{\beta_1}\cos\beta d\beta.$$

Nowe granice całkowania β_1 i β_2 wyznaczają graniczne położenia promieni r poprowadzonych z badanego punktu A do końców rozpatrywanego przewodnika (rys. 21.11). W wyniku całkowania otrzymujemy

$$B = \frac{\mu_0 I}{4\pi r_0}(\sin\beta_1 - \sin\beta_2).$$



Rys. 21.11

W odniesieniu do przewodnika prostoliniowego nieskończenie długiego granice całkowania przyjmą wartości

$$\beta_1 = \frac{\pi}{2} \quad \text{i} \quad \beta_2 = -\frac{\pi}{2},$$

a zatem indukcja B w punkcie A będzie równa

$$B = \frac{\mu_0}{2\pi}\frac{I}{r_0}. \tag{21.16}$$

Jak widać ze wzoru (21.16), indukcja magnetyczna w otoczeniu przewodnika prostoliniowego nieskończenie długiego jest wprost proporcjonalna do natężenia płynącego prądu, a odwrotnie proporcjonalna do odległości badanego punktu od przewodnika z prądem. Wynik otrzymany przy zastosowaniu prawa Biota i Savarta jest oczywiście identyczny z tym, który otrzymaliśmy poprzednio w § 21.4.

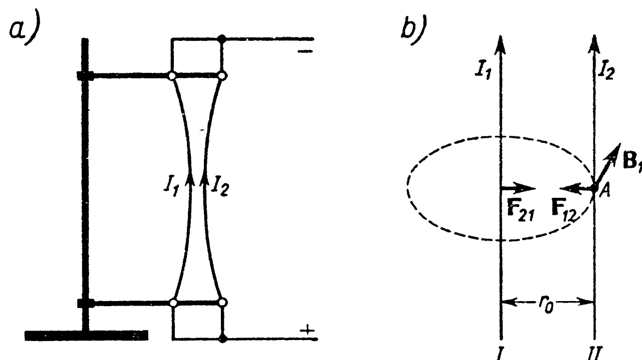
21.6. Wzajemne oddziaływanie przewodów z prądem

Zreasumujmy wnioski z poprzednich paru paragrafów:

1. Pole magnetyczne działa na poruszające się ładunki i na przewody z prądem.
2. W otoczeniu przewodu z prądem powstaje pole magnetyczne.

Zastosujmy te wnioski do dwóch równoległych, nieskończenie długich przewodników rozsuniętych na odległość r_0, w których płyną prądy I_1 i I_2 o kierunkach zgodnych (rys. 21.12a). Doświadczenie wykazuje wzajemne przyciąganie takich przewodów. Jest to zgodne z następującym rozumowaniem: prąd I_1 wytwarza dokoła siebie pole magnetyczne o kołowych liniach indukcji, leżących w płaszczyźnie prostopadłej do przewodnika. Jedna z takich linii, styczna do przewodnika II w punkcie A, zaznaczona jest linią kreskowaną na rys. 21.12b. Kierunek indukcji $\mathbf{B}_1$ w punkcie A styczny do linii indukcji, wskazujący w głąb za płaszczyznę rysunku jest równocześnie prostopadły do kierunku prądu I_2. Na odcinek o długości l przewodnika II, umieszczony w polu magnetycznym pochodzącym od prądu płynącego w przewodniku I, działa zatem siła elektrodynamiczna F_{12} zgodna ze wzorem (21.6):

$$\mathbf{F}_{12} = I_2(\mathbf{l} \times \mathbf{B}_1).$$



Rys. 21.12

Jak łatwo sprawdzić, jest ona skierowana w stronę przewodnika I. Uwzględniając prostopadłość kierunków $\mathbf{l}$ (zgodnego z I_2) i $\mathbf{B}_1$ znajdziemy wartość liczbową siły F_{12} jako

$$F_{12} = I_2 l B_1. \tag{21.17}$$

Ale — zgodnie ze wzorem (21.16) — B_1 w odległości r_0 od nieskończenie długiego przewodnika z prądem I wynosi

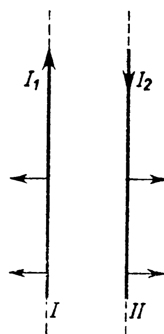
$$B_1 = \frac{\mu_0 I_1}{2\pi r_0}.$$

A zatem po podstawieniu tej wartości do (21.17) otrzymujemy

$$F_{12} = \frac{\mu_0}{2\pi} \frac{I_1 I_2 l}{r_0}. \tag{21.18}$$

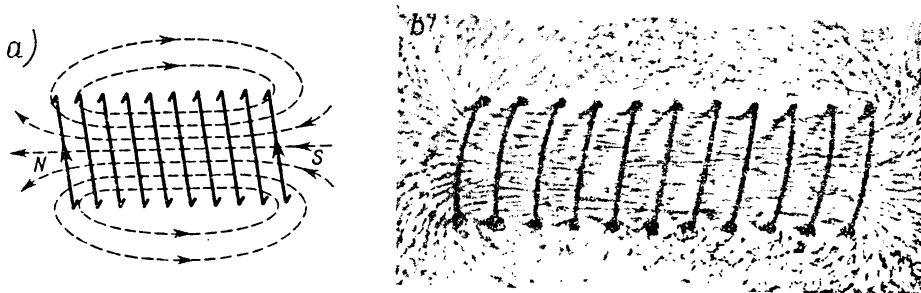
Analogiczne rozumowanie doprowadza do takiego samego wyrażenia na siłę $\mathbf{F}_{21}$ działającą na przewodnik *I*, skierowaną w stronę przewodnika *II*. Mamy więc do czynienia z siłą wzajemnego przyciągania przewodów, w których płyną prądy o kierunkach zgodnych. Czytelnikowi pozostawiamy sprawdzenie, że w przypadku przedstawionym na rys. 21.13 występuje wzajemne odpychanie obu przewodów. Znowu — podobnie jak poprzednio — należy wziąć pod uwagę: 1) pole magnetyczne wytworzone przez prąd płynący w jednym z przewodów, a następnie 2) oddziaływanie tego pola na drugi przewodnik z prądem.

Wzór (21.18) ściśle się wiąże z definicją ampera. Przypominamy podstawowe elementy występujące w definicji: z dwóch nieskończenie długich i cienkich równoległych przewodów, rozsuniętych na odległość $r_0 = 1$ m, wyodrębniamy dwa odcinki, każdy o długości 1 m, mierzymy siłę ich wzajemnego oddziaływania i jeśli równa się ona $2 \cdot 10^{-7}$ N, to mówimy, że każdy z prądów płynących wynosi 1 A (por. ścisłe sformułowanie definicji podane w § 3 „Wiadomości wstępnych"). Podstawiając do wzoru (21.18) odpowiednie wartości liczbowe i jednostki w układzie SI znajdziemy, że rzeczywiście siła F działająca między odcinkami wyżej opisanych przewodów ma wartość $2 \cdot 10^{-7}$ N (wartość przenikalności magnetycznej próżni podana była w § 21.4).



Rys. 21.13

Na zakończenie, nie wnikając w szczegóły rozważań, podamy przybliżone wyniki dotyczące pola magnetycznego solenoidu (rys. 21.14). Można przyjmować, że wewnątrz solenoidu o gęsto nawiniętych zwojach, o średnicy zwoju małej w porównaniu z całkowitą długością l solenoidu, linie indukcji magnetycznej biegną równolegle do osi po-



Rys. 21.14

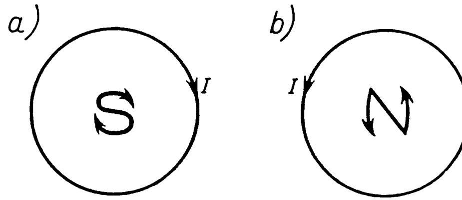
dłużnej solenoidu. Dopiero przy końcach takiego idealnego solenoidu występują odchylenia od równoległości, wykazane na rysunku. Można zatem mówić o jednorodnym polu magnetycznym wewnątrz solenoidu. Indukcja magnetyczna B wewnątrz idealnego solenoidu wyraża się wzorem

$$B = \mu_0 n I, \tag{21.19}$$

gdzie n jest liczbą zwojów na jednostce długości solenoidu. Jak widać ze wzoru, indukcja

B wewnątrz idealnego solenoidu nie zależy ani od jego długości, ani od średnicy zwojów i jest jednakowa we wszystkich punktach wewnątrz solenoidu.

Warto tu może podkreślić, że układ linii indukcji w otoczeniu solenoidu jest analogiczny do układu linii w otoczeniu sztabkowego magnesu. Końcom sztabkowego magnesu przypisujemy nazwy i oznaczenia; biegun północny N i biegun południowy S. Z bieguna N linie indukcji magnetycznej wychodzą w przestrzeń, do bieguna S linie dochodzą z przestrzeni. Analogiczne nazwy można przypisać końcom solenoidu. Łatwo sprawdzić, stosując np. regułę pływaka, że ten koniec solenoidu, w którym — patrząc z zewnątrz — stwierdzimy obieg prądu w kierunku zgodnym z ruchem wskazówek zegara, zachowuje się jak biegun południowy (rys. 21.15a) i odwrotnie — obieg prądu w kierunku przeciwnym do ruchu wskazówek zegara związany jest z biegunem północnym (rys. 21.15b).



Rys. 21.15

Wypisane w specjalny sposób oznaczenia S i N na wspomnianych rysunkach nawiązują do kierunków prądu I w obu obwodach.

Występowanie biegunów magnetycznych N i S rozsuniętych na odległość l kojarzymy z pojęciem dipola magnetycznego (analogia do dipola elektrycznego, por. § 19.2). Właściwości dipolowe w czasie przepływu prądu mają zarówno solenoid, jak i pojedynczy obwód kołowy lub prostokątna ramka z drutu. Obwodom tym przypisujemy *dipolowe momenty magnetyczne*. Obwód z prądem I, którego kontur obejmuje powierzchnię S, ma dipolowy moment magnetyczny $\boldsymbol{\mu}$:

$$\boldsymbol{\mu} = I\mathbf{S}, \tag{21.20}$$

gdzie wektor $\mathbf{S}$ ma kierunek normalnej do powierzchni S, wynikający z obrotu śruby prawoskrętnej w kierunku zgodnym z kierunkiem płynięcia prądu.

Dla obwodu kołowego o promieniu r otrzymamy

$$|\boldsymbol{\mu}| = I\pi r^2, \tag{21.21}$$

a dla obwodu złożonego z n równoległych kołowych zwojów

$$|\boldsymbol{\mu}| = I\pi r^2 n.$$

W rozważaniach dotyczących zachowania się ramki prostokątnej z prądem I w jednorodnym zewnętrznym polu magnetycznym o indukcji $\mathbf{B}$ doszliśmy do wniosku (§ 21.3), że na ramkę działa moment skręcający pary sił, wyrażony wzorem (21.9b).

Obecnie widać, że można ten wzór napisać inaczej, wykorzystując dipolowy moment magnetyczny obwodu:

$$M_{pary} = |\mu| B \sin(\mathbf{n}, \mathbf{B}) \tag{21.22}$$

lub w postaci wektorowej

$$\mathbf{M}_{pary} = \boldsymbol{\mu} \times \mathbf{B}.$$

Momentowi magnetycznemu $\boldsymbol{\mu}$ przypisujemy kierunek wzdłuż osi dipola, a więc prostopadły do powierzchni ramki, a zwrot od bieguna S do bieguna N.

21.7. Natężenie pola magnetycznego

Drugim wektorem stosowanym obok wektora indukcji **B** do opisu pola magnetycznego jest natężenie pola magnetycznego **H**. W celu zdefiniowania tego wektora oprzemy się na wiadomościach z poprzedniego paragrafu, dotyczących jednorodnego pola magnetycznego wewnątrz idealnego, długiego solenoidu. Pole magnetyczne we wnętrzu takiego solenoidu zależy tylko od natężenia I prądu przepływającego przez solenoid i od liczby zwojów n na jednostce długości solenoidu. Można zatem to pole magnetyczne scharakteryzować za pomocą wektora natężenia pola magnetycznego **H**, którego wartość liczbowa zdefiniowana jest jako iloczyn

$$H = In$$

lub

$$H = I\frac{N}{l}, \tag{21.23}$$

gdzie N i l oznaczają odpowiednio ogólną liczbę zwojów solenoidu i jego całkowitą długość.

Wektorowi **H** przypisuje się kierunek zgodny z osią podłużną solenoidu i zwrot wynikający z reguły korkociągu (korkociąg ustawiamy wzdłuż osi podłużnej solenoidu i obracamy jego rączkę w kierunku zgodnym z obiegiem prądu w solenoidzie).

Ze wzoru (21.23) wynika, że *jednostką natężenia pola magnetycznego w układzie SI jest* A/m*.

---

* W dawnym ujęciu, znanym Czytelnikom ze szkoły średniej, natężenie pola magnetycznego kojarzono z siłą działającą na biegun magnetyczny umieszczony w danym punkcie pola. Natężenie pola magnetycznego definiowano jako

$$\mathbf{H} = \frac{\mathbf{F}}{m},$$

gdzie **F** oznacza siłę działającą na biegun magnetyczny o ilości magnetyzmu m, umieszczony w danym punkcie pola. W tym ujęciu natężenie pola magnetycznego wyraża liczbowo siłę działającą na jednostkowy biegun magnetyczny, umieszczony w danym punkcie pola. Ogólnie dawniej stosowaną jednostką natężenia pola był *ersted* (skrót Oe). Związek liczbowy ersteda z jednostką natężenia pola w układzie SI wyraża zależność

$$1\ \text{Oe} = \frac{10^3}{4\pi}\ \text{A/m}.$$

Po wprowadzeniu definicji wektora natężenia pola magnetycznego wzór (21.19) można przepisać w postaci

$$\mathbf{B} = \mu_0 \mathbf{H}. \tag{21.24}$$

Trzeba jednak podkreślić, że ta prosta zależność między **B** i **H** obowiązuje tylko dla próżni (współczynnik μ_0 jest to przenikalność magnetyczna próżni równa $4\pi \times \times 10^{-7}$ kg · m · C^{-2})*. W próżni oba wymienione wektory mają jednakowe kierunki i zwroty i są do siebie wprost proporcjonalne.

Charakter zależności **B** i **H** w różnych ciałach jest różny, w zależności od tego, czy ciało należy do grupy ciał para-, dia-, czy też ferromagnetycznych.

Dla grupy ciał obejmującej *para- i diamagnetyki* (*ciała para- i diamagnetyczne*) obowiązuje zależność:

$$\mathbf{B} = \mu_0 \mu_r \mathbf{H}, \tag{21.25}$$

gdzie μ_r oznacza względną przenikalność magnetyczną danego ośrodka, będącą liczbą oderwaną. Jest to wielkość niezależna od natężenia pola **H**, stała dla danego ośrodka w stałej temperaturze i przy stałej gęstości. Względna przenikalność magnetyczna jest nieco większa od jedynki w przypadku ciał paramagnetycznych, a nieco mniejsza od jedynki w przypadku ciał diamagnetycznych.

Do ciał paramagnetycznych należą m. in. O_2, NO, sole żelaza, niklu i kobaltu, jak również wymienione metale w odpowiednio wysokiej temperaturze (por. niżej). Ciałem paramagnetycznym jest również powietrze. Jego względna przenikalność magnetyczna w warunkach normalnych jest o około $3 \cdot 10^{-7}$ większa od jedynki. Jak widać, ta różnica jest tak niewielka, że w wielu rozważaniach można powietrze traktować jako próżnię.

Do ciał diamagnetycznych należą gazy szlachetne, liczne metale, jak np. Zn, Au, Hg, takie pierwiastki jak krzem i fosfor. Często wymienianym przedstawicielem ciał diamagnetycznych jest bizmut. Jego względna przenikalność magnetyczna różni się od jedynki o $2 \cdot 10^{-4}$.

Odrębną grupę ciał stanowią tzw. *ferromagnetyki* (*ciała ferromagnetyczne*), dla których już nie obserwuje się prostej proporcjonalności między B i H. Do ciał ferromagnetycznych należą Fe, Ni, Co i liczne stopy metaliczne, jak np. permaloj, molipermaloj, supermaloj itd. (zastrzeżenie dotyczące temperatury podane jest niżej).

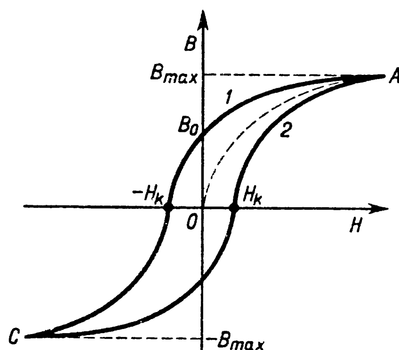
W przypadku ciał ferromagnetycznych względna przenikalność magnetyczna jest liczbą znacznie większą od jedynki (rzędu tysięcy), zmieniającą swą wartość w zależności od **H**, a także w zależności od „poprzednich losów", czyli od tzw. „historii" termicznej i magnetycznej danego ciała. Wpływ historii magnetycznej widać wyraźnie z przebiegu tzw. krzywej histerezy (rys. 21.16).

Krzywą histerezy danego ferromagnetyka otrzymujemy w ten sposób, że na osi odciętych odkładamy natężenia pola, na osi rzędnych — wartości indukcji w danym ferromagnetyku (badany ferromagnetyk jest umieszczony wewnątrz solenoidu, przez który przepuszczamy coraz silniejszy prąd elektryczny, zwiększając w ten sposób **H**

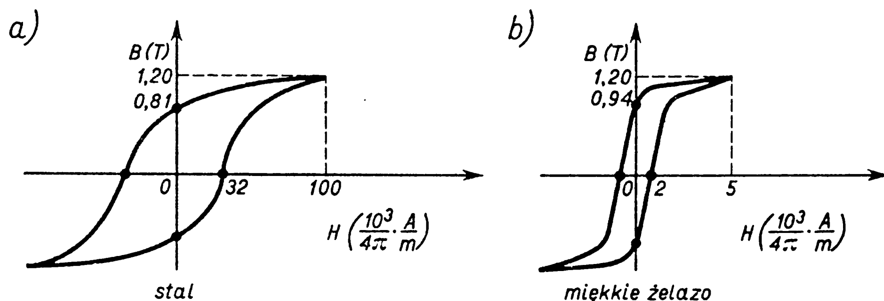
---

* W paragrafie 20.4 wyraziliśmy μ_0 w innych jednostkach. Łatwo wykazać, że oba ujęcia są równoznaczne.

i **B**). Okazuje się, że zależność B od H dla ferromagnetyków ma taki charakter, że biorąc za punkt wyjścia punkt zerowy na wykresie (zupełny brak namagnesowania) przesuwamy się wzdłuż krzywej kreskowanej OA, dochodząc wreszcie do pewnej maksymalnej wartości B_{max}, której już nie można przekroczyć zwiększając H. Jeśli po osiągnięciu tej wartości zaczniemy zmniejszać prąd w solenoidzie wracając stopniowo do wartości $H = 0$, to okaże się, że drogi powrotnej nie odbywamy wzdłuż początkowej krzywej od A do 0, lecz wzdłuż krzywej *1*. Gdy natężenie pola spadnie do zera, indukcja B zachowa wartość B_0, zwaną pozostałością magnetyczną. Zmieniając kierunek płynącego prądu na przeciwny otrzymamy zmianę kierunku **H**. Przy pewnej wartości $H = -H_k$, zwanej koercją, wartość indukcji B spada do zera. Dalszemu wzrostowi ujemnego H towarzyszy stopniowy wzrost ujemnej indukcji B aż do wartości $-B_{max}$. Stosując dalsze zmiany H do wartości zerowej, a następnie do maksymalnej wartości dodatniej wrócimy do punktu A wzdłuż krzywej *2*. Przy zastosowaniu dostatecznie szerokich granic zmienności $+H$ i $-H$ otrzymamy właśnie przedstawioną na wykresie krzywą zamkniętą, zwaną pętlą histerezy.



Rys. 21.16



Rys. 21.17

Przebieg pętli może być bardzo różny dla różnych ferromagnetyków, do których zaliczamy żelazo, nikiel, kobalt oraz liczne stopy metali wymienionych, jak też i innych. Przykładowe rysunki 21.17a i b przedstawiają przebieg krzywych histerezy dla stali i miękkiego żelaza. Badania wykazały, że pole pętli histerezy jest proporcjonalne

do energii rozpraszanej w postaci ciepła w czasie przemagnesowywania się ferromagnetyka. A zatem, spoglądając na rys. 21.17 stwierdzimy, że miękkie żelazo (i stopy o podobnym charakterze pętli histerezy) będzie odpowiednim materiałem na rdzenie w obwodach prądu zmiennego (małe straty energii w czasie przemagnesowywania). Natomiast materiałem odpowiednim na sztuczne magnesy trwałe będą m. in. różne rodzaje stali (stale chromowe, kobaltowe, wolframowe i in.).

Przebieg pętli histerezy danego materiału zmienia się w zależności od temperatury: ze wzrostem temperatury obniża się przebieg krzywych histerezy: maleje wartość B_{max}.

Każdemu z ferromagnetyków odpowiada określona temperatura, zwana *temperaturą Curie*, w której w sposób gwałtowny traci on swoje właściwości ferromagnetyczne, przekształcając się w ciało paramagnetyczne. Temperatury Curie równają się odpowiednio: dla żelaza 1043K, dla kobaltu 1400K i dla niklu 631K.

21.8. Wektor namagnesowania

Przy omawianiu magnetycznych właściwości różnych substancji wygodnie jest posługiwać się tzw. *wektorem namagnesowania* **M**. Zrozumienie definicji tej nowej wielkości wymaga kilku uwag wprowadzających.

Jak już wspominaliśmy, właściwości magnetyczne ciała są efektem wtórnym, wynikającym z ruchu ładunków elektrycznych. Właściwości magnetyczne należy zatem przypisywać oddzielnym atomom (cząsteczkom), w ich skład bowiem wchodzą elektrony odbywające określone ruchy.

Z ruchem elektronu po orbicie dookoła jądra można kojarzyć nie tylko istnienie *orbitalnego momentu pędu*, lecz także istnienie dipola magnetycznego o pewnym momencie magnetycznym, zwanym *momentem magnetycznym orbitalnym*, którego wartość zależy od częstotliwości krążenia elektronu i pola objętego jego orbitą (por. wzór (21.21)).

Z ruchem elektronu dookoła własnej osi można kojarzyć nie tylko *moment pędu zwany spinem elektronu*, lecz także *magnetyczny moment spinowy*.

W zależności od struktury atomu (cząsteczki) wypadkowy moment magnetyczny (orbitalny i spinowy) może być równy zeru lub różny od zera.

Biorąc pod uwagę nie pojedyncze atomy lub cząsteczki, lecz skończone objętości badanych substancji, można mówić o sumie wektorowej magnetycznych momentów atomów lub cząsteczek, czyli o wypadkowym momencie magnetycznym, przypadającym na daną objętość. Wypadkowy moment magnetyczny, przypadający na jednostkę objętości, stanowi tzw. *wektor namagnesowania* **M**:

$$\mathbf{M} = \lim_{V \to 0} \frac{\sum_{i=1}^{N} \boldsymbol{\mu}_i}{V}, \tag{21.26}$$

gdzie N oznacza liczbę atomów (cząsteczek) w objętości V substancji, $\boldsymbol{\mu}_i$ — moment magnetyczny i-tego atomu (cząsteczki). Łatwo sprawdzić, że jednostką namagnesowania **M** w układzie SI jest A/m, a więc jednostka taka sama, jak w przypadku **H**.

Powiedzieliśmy wyżej, że atomy (cząsteczki) danej substancji w zależności od swej struktury mogą mieć wypadkowy moment magnetyczny równy zeru lub różny od zera. Tę właściwość można przyjąć za podstawę podziału ciał na dia- i paramagnetyki.

Atomy (cząsteczki) ciał diamagnetycznych mają — przy braku zewnętrznego pola — wypadkowy moment magnetyczny równy zeru.

Atomy (cząsteczki) ciał paramagnetycznych mają — przy braku zewnętrznego pola — wypadkowy moment magnetyczny nierówny zeru. W odniesieniu jednak do skończonych elementów objętości badanej substancji wypadkowy moment magnetyczny przyjmuje wartość zerową dzięki całkowicie bezładnemu przestrzennemu rozkładowi atomowych (cząsteczkowych) momentów magnetycznych.

Zastosowanie zewnętrznego pola magnetycznego narastającego od zera do pewnej wartości H wywołuje w obwodzie orbitalnego prądu elektronowego dodatkowy prąd indukcyjny (por. rozdział następny) o takim kierunku (reguła Lenza), że jego pole magnetyczne jest skierowane przeciwnie do kierunku zewnętrznego pola*. *Powstanie tego indukowanego pola powoduje osłabienie strumienia indukcji charakterystyczne dla diamagnetyzmu.*

Efekt diamagnetyczny występuje w zasadzie we wszystkich ciałach, gdyż ruch orbitalny elektronów odbywa się we wszystkich atomach. Jest to jednak efekt stosunkowo słaby i daje się zaobserwować tylko wtedy, gdy nie jest maskowany silniejszym efektem paramagnetycznym.

Zastosowanie zewnętrznego pola magnetycznego w odniesieniu do ciał paramagnetycznych wywołuje obok wyżej opisanego słabego efektu diamagnetycznego silniejszy *efekt paramagnetyczny* związany z tym, że tym razem każdy atom (cząsteczka) ma moment magnetyczny nierówny zeru, a więc zachowuje się jak mały dipol magnetyczny. Pod wpływem zewnętrznego pola wspomniane dipole usiłują obrócić się tak, by zająć położenie zgodne z kierunkiem pola zewnętrznego. Chaotyczne ruchy cieplne atomów (cząsteczek) utrudniają to uporządkowanie (wyraźny wpływ temperatury). W wyniku działania tych dwóch przeciwstawnych czynników ustala się pewien stan równowagi. Uzyskane *częściowe uporządkowanie ustawienia dipoli magnetycznych powoduje powiększenie strumienia indukcji, charakterystyczne dla zjawiska paramagnetyzmu.*

W ścisłym związku z omówionymi zjawiskami pozostaje wciąganie ciał paramagnetycznych i wypychanie ciał diamagnetycznych z obszaru pola w czasie magnesowania tych ciał.

Właściwości ferromagnetyków wiążą się z określoną strukturą krystaliczną i z silnymi polami sił wewnętrznych, istniejącymi w tych ciałach. Zakłada się, że spinowe momenty magnetyczne elektronów atomów znajdujących się w węzłach sieci krystalicznej silnie na siebie oddziaływają. W wyniku tych oddziaływań w pewnych niewielkich obszarach ferromagnetyka, zwanych *domenami*, pojawia się pewne uporządkowanie, polegające na równoległym ustawieniu spinowych momentów magnetycznych atomów sąsiadujących ze sobą w sieci krystalicznej. Przy braku pola zewnętrznego w sąsiednich domenach

* Ogólnie biorąc, gdy pole zewnętrzne jest ukośne względem płaszczyzny orbity, to powstaje ruch precesyjny osi orbity dookoła kierunku pola. Ta precesja wywołuje powstanie dodatkowego momentu magnetycznego skierowanego przeciwnie do kierunku **H** zewnętrznego.

ustawienie spinów jest różne. Po zastosowaniu rosnącego, stosunkowo słabego pola zewnętrznego, następuje powiększanie się obszaru tych domen, w których namagnesowanie **M** ma kierunek najbardziej zbliżony do kierunku pola zewnętrznego, kosztem domen sąsiednich. Po osiągnięciu dość silnych pól następuje skokowa zmiana kierunku namagnesowania poszczególnych domen lub kryształu jako całości. Proces namagnesowania ferromagnetyków ma zatem charakter skokowy, co udało się stwierdzić doświadczalnie (tzw. efekt *Barkhausena*).

Jak już wspominaliśmy, za właściwości ferromagnetyczne odpowiedzialne są momenty magnetyczne spinowe i struktura sieci krystalicznej. W temperaturze Curie te właściwości zanikają, gdyż w tej temperaturze zachodzi zmiana struktury sieci krystalicznej. W tej temperaturze ulegają też zmianie liczne współczynniki charakteryzujące daną substancję.

Pytania i zadania

1. Co jest źródłem pola magnetycznego?

2. Jakie oddziaływania wykorzystano do zdefiniowania wektora indukcji magnetycznej **B**?

3. W jakich jednostkach wyrażamy indukcję magnetyczną?

4. Jakim wzorem wyraża się siła oddziaływania pola magnetycznego na poruszający się ładunek? Jaka reguła kierunkowa określa kierunek tej siły?

5. Co nazywamy siłą Lorentza?

6. Pamiętając o tym, że kierunek stałej siły, działającej w stałym polu magnetycznym na przelatujący poprzecznie ($\mathbf{v} \perp \mathbf{B}$) elektron, jest stale prostopadły do **v**, określić charakter toru ruchu elektronu. Czy z ruchem elektronu po tym torze wiąże się zmiana energii kinetycznej?

7. Jaka jest różnica skutków działania sił pochodzących od pola elektrycznego i pola magnetycznego w przypadku poprzecznego przelotu elektronu przez te pola?

8. Podać określenie strumienia indukcji i jego jednostki.

9. Podać wzory wyrażające siłę elektrodynamiczną, działającą podczas przepływu prądu I na: a) odcinek dl przewodu umieszczony w polu magnetycznym, b) odcinek l przewodu prostoliniowego umieszczonego w jednorodnym polu magnetycznym.

10. Jakie są skutki działania sił elektrodynamicznych na boki prostokątnej cewki swobodnie zawieszonej w jednorodnym polu magnetycznym, którego indukcja **B** skierowana jest a) równolegle, b) prostopadle do płaszczyzny cewki. Uzasadnić szczegółowo odpowiedź.

11. Podać charakterystykę opisową rozkładu linii indukcji w otoczeniu przewodników: a) liniowych, b) kołowych, przez które płynie prąd elektryczny.

12. Jakie rozumowanie doprowadza do wzoru na siłę wzajemnego oddziaływania przewodników, przez które płyną prądy elektryczne? Jaki wzór wyraża siłę działającą na odcinek l długiego przewodnika, przez który płynie prąd I_1, jeżeli w odległości r od niego znajduje się drugi, równoległy do niego, długi przewodnik, przez który płynie prąd I_2?

13. Wypisać dokładną definicję ampera. Sprawdzić, czy jest ona zgodna z definicją podaną w podręczniku.

14. Jak można scharakteryzować pole magnetyczne, powstające wewnątrz długiego solenoidu o promieniu małym w stosunku do długości?

15. Co rozumiemy przez moment dipolowy magnetyczny? Jakie czynniki decydują o momencie magnetycznym obwodu, przez który płynie prąd elektryczny?

16. Obliczyć indukcję magnetyczną w odległości $r = 10\,\text{cm}$ od długiego prostoliniowego przewodnika, przez który płynie prąd $I = 2A$?

Odp. $4 \cdot 10^{-6}$ T.

17. Indukcja magnetyczna w środku obwodu kołowego o promieniu $R = 10$ cm równa się $2\pi \cdot 10^{-5}$ T. Jaki prąd płynie przez ten obwód?

Odp. 10 A.

18. Obliczyć indukcję magnetyczną w środku obwodu w kształcie kwadratu o boku $a = 8$ cm, jeżeli natężenie prądu w tym obwodzie równa się 10 A.

Odp. $10^{-4} \cdot \sqrt{2}$ T.

19. Obliczyć siłę działającą na elektron przebiegający przez jednorodne pole magnetyczne o indukcji $B = 3$ T, skierowanej pionowo w dół. Energia kinetyczna elektronu wynosi $0,5 \cdot 10^{-17}$ J. Prędkość jego skierowana jest poziomo w stronę północy. Masa elektronu $m_e = 9 \cdot 10^{-31}$ kg.

Odp. $1,6 \cdot 10^{-12}$ N, wschód.

20. W długim, pionowym przewodzie płynie z góry w dół prąd elektryczny $I = 6$ A. W jakiej odległości od tego przewodu wypadkowa indukcja, wynikająca z istnienia pola magnetycznego prądu i pola magnetycznego Ziemi, będzie miała kierunek pionowy. Składowa pozioma indukcji magnetycznej Ziemi $B_h = 0,2 \cdot 10^{-4}$ T.

Odp. 6 cm, po stronie wschodniej.

21. W jednorodnym polu magnetycznym o indukcji $B = 0,2$ T umieszczono prostokątną cewkę o bokach $a = 6$ cm, $b = 4$ cm. Jak wielki jest moment pary sił działających na cewkę? Cewka składa się z $n = 25$ zwojów i płynie przez nią prąd $I = 2$ A. Wektor indukcji skierowany jest równolegle do krótszego boku cewki.

Odp. $24 \cdot 10^{-3}$ N · m.

22. Okres drgań małej igiełki magnetycznej w polu magnetycznym Ziemi równa się 0,7 s, zaś okres jej drgań w polu magnetycznym wewnątrz długiego solenoidu, przez który płynie prąd, równa się 0,1 s. Zaniedbując tłumienie drgań i przyjmując, że składowa pozioma indukcji magnetycznej Ziemi $B_h = 0,18 \cdot 10^{-4}$ T, wyznaczyć indukcję B wewnątrz solenoidu.

Odp. $8,8 \cdot 10^{-4}$ T.

23. Obliczyć indukcję magnetyczną wewnątrz solenoidu o długości $l = 80$ cm i liczbie zwojów $n = 1600$, jeżeli przez solenoid płynie prąd $I = 2$ *A*. Zakładamy, że promień solenoidu jest mały w porównaniu z jego długością.

Odp. $16 \cdot 10^{-4}$ T.

24. Przez dwa długie, prostoliniowe, równoległe do siebie przewodniki umieszczone w odległości $d = 50$ cm płyną prądy o jednakowym natężeniu $I = 10$ A w zgodnych kierunkach. Jak wielka i jak skierowana siła działa na odcinek o długości $l = 5$ m każdego z przewodników?

Odp. $2 \cdot 10^{-4}$ N.

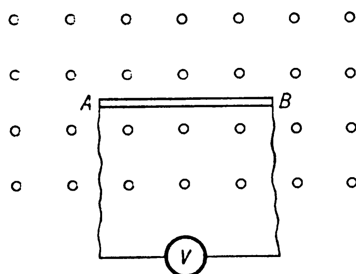
ROZDZIAŁ 22

Indukcja elektromagnetyczna. Prądy zmienne

22.1. Uwagi wstępne

Odkrycie zjawiska indukcji elektromagnetycznej związane z nazwiskiem Faradaya (1831 r.) ma bardzo istotne znaczenie dla współczesnej elektrotechniki. Na zjawisku tym bowiem opiera się m. in. działanie podstawowych źródeł energii elektrycznej.

Omówienie tego zjawiska rozpoczniemy od zestawienia szeregu faktów doświadczalnych i wypływających z nich wniosków. Następnie przejdziemy do wyjaśnienia zjawisk na gruncie teorii elektronowej i do analizy ich strony ilościowej.



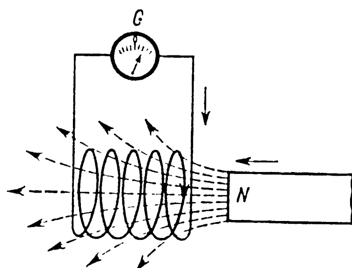
Rys. 22.1

Podstawowe dla zjawiska indukcji elektromagnetycznej doświadczenie przedstawione jest na rys. 22.1. Przewodnik AB znajduje się w polu magnetycznym o indukcji $\mathbf{B}$, prostopadłej do płaszczyzny rysunku, skierowanej przed płaszczyznę rysunku (kropki na rysunku). Gdy przewodnik ten gwałtownym ruchem usuwamy z pola magnetycznego, woltomierz wysokoomowy wykazuje powstanie krótkotrwałego napięcia między punktami A i B, zwanego *siłą elektromotoryczną indukowaną* (*SEM indukowana*). Wsuwanie tego przewodnika do pola magnetycznego wywołuje powstanie SEM o przeciwnym znaku. Przesuwanie przewodnika wzdłuż linii indukcji nie daje żadnego wychylenia wskazówki miernika. A zatem: *przecinanie linii indukcji przez przewodnik*

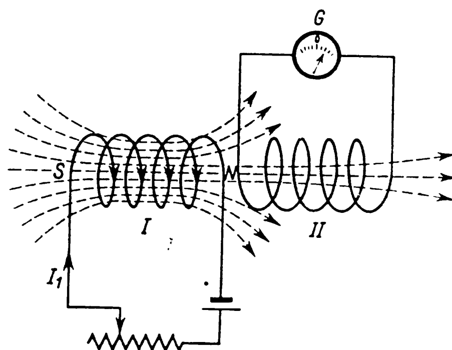
podczas jego ruchu przez pole magnetyczne powoduje pojawienie się na końcach przewodnika siły elektromotorycznej indukowanej.

Jeśli przewodnik użyty do doświadczenia ma postać zamkniętego obwodu z włączonym miernikiem prądu, to każda zmiana strumienia magnetycznego przechodzącego przez powierzchnię objętą konturem obwodu wywołuje wychylenie wskazówki miernika. Świadczy to o przepływie w obwodzie zamkniętym prądu, który nazywamy *prądem indukowanym.*

Wymienimy kilka możliwych sposobów wywoływania prądów indukowanych w obwodach zamkniętych, czyli, innymi słowy, kilka sposobów wywoływania zmiennego w czasie strumienia magnetycznego Φ, przechodzącego przez powierzchnię objętą konturem obwodu. A więc przede wszystkim: 1) ruch magnesu względem obwodu lub odwrotnie, gdyż istotne znaczenie ma tutaj ruch względny magnesu i obwodu, pociągający za sobą zmianę opisanego wyżej strumienia magnetycznego, 2) ruch obwodu z prądem względem drugiego obwodu zamkniętego, 3) zmiany natężenia prądu w jednym obwodzie, wywołujące zmiany strumienia magnetycznego w odpowiednio umieszczonym obwodzie drugim.



Rys. 22.2



Rys. 22.3

Jako ilustracja sposobu pierwszego może służyć doświadczenie przedstawione na rys. 22.2. Szybkim ruchem wprowadzamy biegun N do wnętrza cewki, której końce dołączone są do galwanometru *G*. Towarzyszy temu chwilowe wychylenie wskazówki w jedną stronę. Usuwając tenże biegun wywołamy wychylenie jej w stronę przeciwną. Zmiana znaku wprowadzanego bieguna magnetycznego również zmienia kierunek wychylenia wskazówki przyrządu. Przyczyną wystąpienia prądu indukowanego w obwodzie jest zmiana strumienia magnetycznego przechodzącego przez obwód cewki, towarzysząca ruchowi magnesu. (W warunkach statycznych, gdy magnes nie zmienia swego położenia względem cewki, wskazówka przyrządu stoi na zerze: prąd w obwodzie nie płynie.) Analizując dalej powstający w obwodzie prąd ustalamy, jaki jest jego kierunek. Mówi o tym *reguła Lenza*: kierunek powstającego prądu indukowanego musi być taki, aby wywołane przezeń pole magnetyczne przeciwdziałało tym zmianom, które wywołały jego powstanie. A zatem kierunek prądu indukowanego w cewce musi być

taki, aby towarzyszące mu pole magnetyczne utrudniało ruch magnesu. Jeśli zatem do zwojnicy zbliżamy biegun N, to musi on podlegać odpychaniu, czyli w uzwojeniu oglądanym od strony zbliżanego bieguna musi powstać prąd o kierunku przeciwnym do ruchu wskazówek zegara. W podobny sposób można przeanalizować kierunek prądu indukowanego, wywołanego zbliżaniem bieguna S lub oddalaniem N lub S.

Rozważanie drugiego sposobu wzbudzania prądu indukowanego nie wnosi w zasadzie nic nowego. Pamiętamy bowiem, że obwód z prądem zachowuje się jak dipol magnetyczny o tak rozmieszczonych biegunach, jak to wynika z kierunku płynięcia prądu (por. rys. 21.15). Przypadek drugi sprowadza się zatem do pierwszego.

W przypadku trzecim należy pamiętać o tym, że strumień magnetyczny wzbudzany wskutek przepływu prądu elektrycznego jest funkcją natężenia tego prądu. Wyobraźmy sobie dwa obwody: jeden ze źródłem prądu i np. suwakowym opornikiem, drugi — połączony z galwanometrem G (rys. 22.3). Oba obwody są ustawione tak, że linie indukcji magnetycznej **B** wywołane przepływem prądu w obwodzie pierwszym przecinają powierzchnię objętą konturem obwodu drugiego. Każdej zmianie prądu, a więc również każdemu otwarciu i zamknięciu obwodu *I*, towarzyszy krótkotrwałe wychylenie wskazówki miernika w obwodzie *II*. Oddziaływanie indukcyjne dwóch obwodów elektrycznych obejmujemy ogólną nazwą *indukcji wzajemnej*. Omówimy ją szczegółowiej w § 22.4.

Objawy zjawiska indukcji elektromagnetycznej można też obserwować w odosobnionym obwodzie elektrycznym, jeśli zachodzą w nim zmiany prądu (a więc np. przy zamykaniu lub otwieraniu obwodu). Zmiany te wywołują zmiany strumienia magnetycznego przechodzącego przez dany obwód, a te pociągają za sobą powstanie prądu indukowanego. To zjawisko, zwane *indukcją własną*, rozważamy w § 22.5.

22.2. Prawo indukcji elektromagnetycznej Faradaya

Wartość SEM indukowanej otrzymujemy z następujących rozważań prowadzących do prawa Faradaya. Utwórzmy obwód w kształcie prostokątnej ramki, której bok AK (rys. 22.4) stanowi ruchomą poprzeczkę mogącą się ślizgać bez tarcia wzdłuż boków CD i EG. Do obwodu załączamy tylko miernik prądu. Ramkę umieszczamy w jednorodnym polu magnetycznym o indukcji **B**, prostopadłej do płaszczyzny ramki, skierowanej np. za płaszczyznę rysunku, co zaznaczamy wypełniając obszar pola magnetycznego krzyżykami. Siłą zewnętrzną przesuwamy AK z położenia *I* do położenia *II* o odległość ds. Ruchowi temu towarzyszy zmiana strumienia magnetycznego obejmowanego konturem obwodu oraz powstanie prądu indukowanego. Kierunek tego prądu musi być taki, aby powstająca siła elektrodynamiczna **F** przeciwdziałała zewnętrznej sile przesuwającej poprzeczkę. Jeżeli przesuwanie poprzeczki AK odbywa się ruchem jednostajnym, to $\mathbf{F}_{zewn} = -\mathbf{F} = -I(\mathbf{l}\times\mathbf{B})$. Ponieważ $\sin(\mathbf{l}, \mathbf{B}) = 1$, więc pracę dW siły zewnętrznej na drodze ds można wyrazić wzorem

$$dW = -IlB\,ds.$$

Iloczyn $l\,ds = dS$ wyraża zmianę powierzchni objętej konturem obwodu. A zatem

$$B\,dS = d\Phi, \tag{22.1}$$

$$dW = -I d\Phi.$$

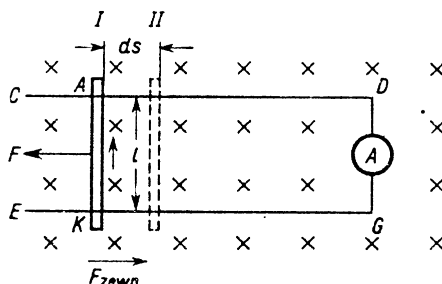
Kosztem tej pracy powstaje energia prądu indukowanego, równa iloczynowi SEM indukowanej E_{ind}, natężenia prądu indukowanego I i czasu jego przepływu dt:

$$-I d\Phi = E_{ind} I dt,$$

skąd

$$E_{ind} = -\frac{d\Phi}{dt}. \tag{22.2}$$

Wzór ten wyraża *prawo indukcji elektromagnetycznej Faradaya. SEM indukowana jest proporcjonalna do szybkości zmiany strumienia magnetycznego w danym obwodzie.* Znak minus we wzorze nawiązuje do reguły kierunkowej Lenza i przypomina, że *kierunek prądu indukowanego jest zawsze taki, że pole magnetyczne przezeń wywołane przeciwstawia się zmianie strumienia magnetycznego zewnętrznego.*



Rys. 22.4

Wzór (22.2) wyprowadziliśmy w odniesieniu do prostokątnej, pojedynczej ramki. W przypadku cewki złożonej z n zwojów izolowanego drutu, położonych blisko siebie, siły elektromotoryczne indukowane w poszczególnych zwojach dodają się i wypadkowa SEM indukowana równa się

$$E_{ind} = -n\frac{d\Phi}{dt}. \tag{22.3}$$

Wzór wyrażający SEM indukowaną można też wyrazić w postaci wskazującej na bezpośrednią zależność SEM od prędkości ruchu przewodnika w jednorodnym polu magnetycznym o indukcji **B**. Uwzględniając (22.2) i (22.1) znajdujemy

$$E_{ind} = -\frac{B dS}{dt} = -Bl\frac{ds}{dt} = -Blv, \tag{22.4}$$

gdzie $v = ds/dt$ jest prędkością, z jaką przewodnik o długości l jest przesuwany w polu magnetycznym prostopadle do kierunku indukcji **B**. Kierunek prądu indukowanego wynika bezpośrednio z pomocniczej *reguły kierunkowej trzech palców prawej ręki*, przedstawionej na rys. 22.5.

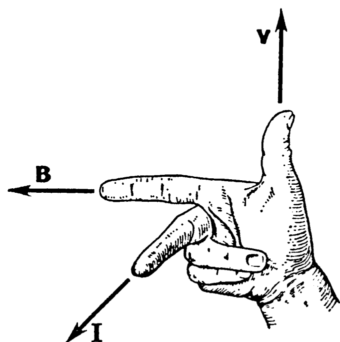
Podczas przesuwania przewodnika w kierunku tworzącym kąt α z indukcją **B** zamiast wzoru (22.4) mielibyśmy

$$E_{ind} = -Blv\sin\alpha, \quad \text{czyli} \quad E_{ind} = -Blv\sin(\mathbf{v}, \mathbf{B}). \tag{22.5}$$

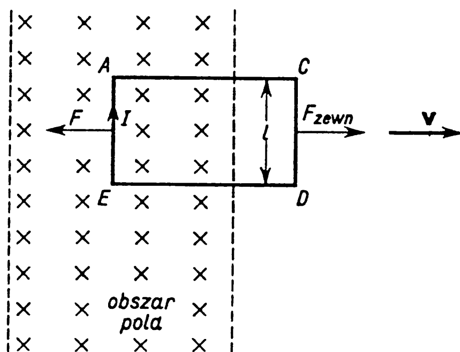
Analogiczne wyrażenia otrzymalibyśmy rozważając usuwanie lub wprowadzanie do pola o stałej indukcji **B** obwodu prostokątnego *ACDE* o powierzchni prostopadłej do **B** (rys. 22.6).

Warto podkreślić, że SEM indukowana nie zależy od rodzaju przewodnika, w jakim jest indukowana. Opór elektryczny R przewodnika ma natomiast wpływ na wartość prądu, a mianowicie

$$I = \frac{E_{\text{ind}}}{R} = \frac{1}{R}\frac{d\Phi}{dt} \quad \text{lub} \quad I = \frac{Blv}{R}. \tag{22.6}$$



Rys. 22.5



Rys. 22.6

Na podstawie wzoru (22.6) można też napisać wyrażenie na moc P prądu powstającego w rozważanym przypadku:

$$P = I^2R = \frac{B^2l^2v^2}{R}.$$

W rozpatrzonym zjawisku mamy przykład przemian energetycznych: kosztem energii mechanicznej związanej z pracą siły zewnętrznej powstaje energia elektryczna, która następnie przekształca się na ciepło Joule'a.

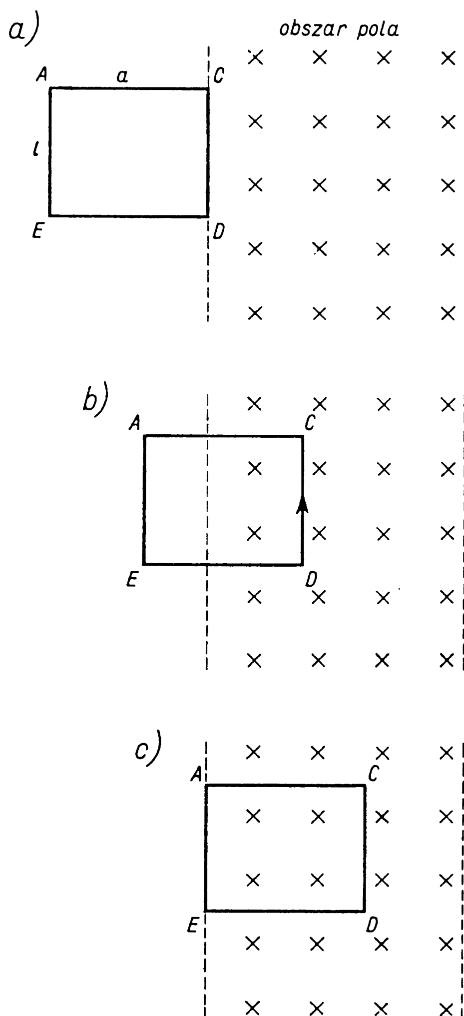
Rozpatrzmy bardziej szczegółowo SEM indukowaną w ramce prostokątnej (rys. 22.7) o szerokości l i o długości a, w czasie jej ruchu ze stałą prędkością **v** w obszarze jednorodnego pola magnetycznego o indukcji **B**, skierowanej prostopadle za płaszczyznę rysunku.

Rozważmy kolejno następujące etapy ruchu ramki:

1. Ramka znajduje się w całości poza obrębem pola. Strumień magnetyczny Φ równa się zeru aż do chwili, gdy bok ramki *CD* pokryje się z brzegiem obszaru pola (rys. 22.7a). Do tej chwili SEM indukowana równa się zeru.

2. Wskutek dalszego ruchu ramki w prawo narasta w ramce strumień Φ proporcjonalnie do drogi przebytej przez *CD* w polu magnetycznym (rys. 22.7b). Wobec stałej prędkości ruchu ramki zmiana strumienia w każdej jednostce czasu (równa $d\Phi/dt$) jest stała, a zatem w obwodzie indukuje się SEM o stałej wartości. Łatwo sprawdzić, że wywołuje ona w ramce prąd o kierunku przeciwnym do ruchu wskazówek zegara.

3. Zmiana strumienia magnetycznego opisana w punkcie 2 zachodzi w obwodzie tylko do chwili, gdy cała ramka znajdzie się w obrębie pola (rys. 22.7c) (tzn. bok *AE* pokryje się z lewą granicą obszaru pola). Dalszy ruch — aż do chwili, gdy bok *CD* osiągnie prawy brzeg pola — nie powoduje zmiany strumienia magnetycznego, a zatem nie wywołuje SEM indukowanej.



Rys. 22.7

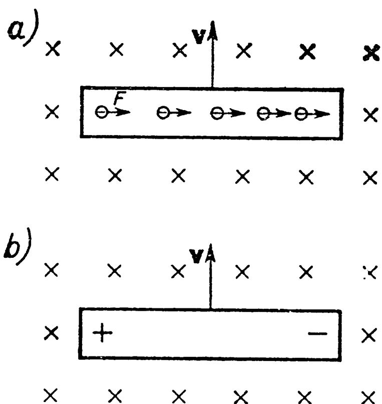
4. Z chwilą gdy bok *CD* wyminie prawą granicę obszaru pola magnetycznego, rozpoczyna się znów zmiana strumienia magnetycznego (tym razem strumień maleje). Towarzyszy temu pojawienie się SEM indukowanej w obwodzie ramki. Czytelnik sam odpowie na pytania: a) czy SEM w tym przypadku będzie stała? b) w jakim kierunku będzie w ramce płynął prąd indukowany?

5. Ramka wysunęła się całkowicie z obszaru pola. Dalszy jej ruch nie pociąga za sobą ani zmiany strumienia magnetycznego, ani powstania SEM indukowanej.

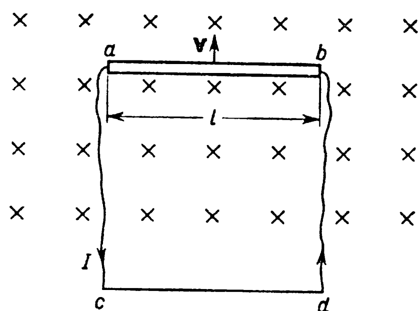
Wyjaśnienie powstawania SEM indukowanej na gruncie teorii elektronowej metali przedstawia się bardzo prosto. W każdym odcinku l przewodnika metalowego występują swobodne elektrony. Wprawiając ten odcinek w ruch z prędkością $\mathbf{v}$ w polu magnetycznym spowodujemy wystąpienie znanej już nam siły, działającej na poruszające się ładunki, równej

$$\mathbf{F} = -e(\mathbf{v} \times \mathbf{B}).$$

Pod wpływem tej siły (rys. 22.8a) swobodne elektrony przesuną się w kierunku jednego z końców przewodnika l. Na tym końcu wystąpi nadmiar ładunków ujemnych, zaś na końcu przeciwległym ich niedobór, tzn. pojawią się ładunki dodatnie (rys. 22.8b). Między końcami przewodnika powstanie różnica potencjałów stanowiąca właśnie SEM indukowaną. Kierunek przypisywany SEM indukowanej jest od — do + (podobnie jak kierunek SEM źródła prądu stałego: ogniwa, akumulatora itp. — wewnątrz źródła od bieguna ujemnego do bieguna dodatniego).



Rys. 22.8



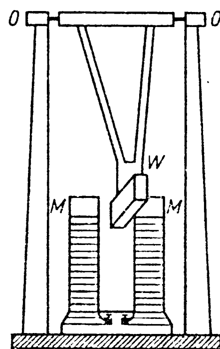
Rys. 22.9

Połączmy końce rozważanego odcinka l przewodnikiem *acdb* (rys. 22.9) tak, by część *cd* obwodu zamkniętego była poza polem magnetycznym. W tym przypadku podczas ruchu *ab* coraz nowe swobodne elektrony napływające do tego odcinka z reszty obwodu będą bez zahamowania na końcach *a* i *b* przesuwane coraz dalej. W zamkniętym obwodzie odbywać się będzie ruch elektronów (równoważny przepływowi prądu elektrycznego) tak długo, jak długo z ruchem *ab* przez pole magnetyczne związana będzie zmiana strumienia magnetycznego w obwodzie *abdca*. Kierunek prądu indukowanego (odniesiony do ruchu ładunków dodatnich) jest taki sam, jak kierunek SEM indukowanej na odcinku *ab*. Wygodniej jest oczywiście kierunek tego prądu określać posługując się regułą trzech palców prawej ręki. Raz jeden jednak warto się przekonać, że oba sposoby prowadzą do tego samego wyniku.

Pytanie. Wiadomo, że w dielektrykach nie ma swobodnych elektronów. Czy ruch pręta z dielektryka w obszarze pola magnetycznego jednorodnego w kierunku prostopadłym do indukcji **B** wywoła jakieś zmiany w dielektryku?

22.3. Prądy Foucaulta

Prądy indukowane występują nie tylko w obwodach elektrycznych z cienkich przewodników, jeśli obejmują one zmienny w czasie strumień magnetyczny, lecz także pojawiają się w masywnych przewodnikach (bryłach, prętach itp.), umieszczonych w zmiennym polu magnetycznym, jako *indukowane prądy wirowe* zwane *prądami Foucaulta.* Kierunek tych prądów podlega również regule Lenza. Jeżeli więc zmienność strumienia magnetycznego wynika np. z ruchu masywnej bryły w polu magnetycznym (por. znane z kursu szkolnego wahadło Foucaulta, przedstawione na rys. 22.10), to ruch ten zostaje gwałtownie zahamowany. Efekt jest taki, jakby bryła wpadała w gęsty, lepki ośrodek.



Rys. 22.10

Podobnie blaszka metalowa nieferromagnetyczna puszczona swobodnie między biegunami elektromagnesu (**B** poziome) spada bardzo powoli, dopóki nie wydostanie się z obszaru pola magnetycznego.

Wobec małych elektrycznych oporów właściwych metali prądy wirowe w bryłach, prętach osiągają duże natężenia, powodując wydzielanie dużych ilości ciepła, a tym samym duże straty energii. Celem zmniejszenia tych strat — szczególnie dużych przy szybkich zmianach Φ — zastępuje się masywne przewodniki oddzielnymi blachami lub cienkimi prętami umieszczonymi prostopadle do kierunku płaszczyzn spodziewanych prądów Foucaulta. Tak np. rdzenie transformatorów buduje się z oddzielnych blach ze specjalnego gatunku stali o zwiększonym oporze elektrycznym właściwym dzięki wprowadzeniu krzemu.

Prądy Foucaulta mogą być jednak także wykorzystywane jako zjawisko użyteczne, np. do elektrycznego hamowania, do topienia metali itp.

22.4. Indukcja wzajemna dwóch obwodów

Zajmiemy się obecnie nieco szczegółowiej prądami indukowanymi, wynikającymi z wzajemnego oddziaływania na siebie dwóch obwodów elektrycznych. Niech to będą np. obwody jak na rys. 22.3. Źródło prądu stałego wywołuje w cewce pierwszej przepływ prądu elektrycznego I_1, wskutek czego powstaje strumień magnetyczny Φ_1, przechodzący przez tę cewkę. Część tego strumienia, a mianowicie strumień Φ_{21}, proporcjonalny do I_1, przechodzi przez cewkę drugą:

$$\Phi_{21} = L_{21} I_1, \tag{22.7}$$

gdzie współczynnik proporcjonalności L_{21} jest tzw. *indukcyjnością wzajemną* (*współczynnikiem indukcji wzajemnej*) obu obwodów, zależną od ich ustawienia względem siebie,

od ich kształtów, rozmiarów, przenikalności magnetycznej środowiska itp. (zakładamy, że środowisko nie ma właściwości ferromagnetycznych). *Indukcyjność wzajemna dwóch obwodów* wyraża liczbowo strumień magnetyczny przechodzący przez obwód drugi przy przepływie prądu 1 A w obwodzie pierwszym.

W czasie pojawiania się strumienia magnetycznego w cewce drugiej powstaje w niej SEM indukowana równa

$$E_2 = -\frac{d\Phi_{21}}{dt},$$

zgodnie z prawem indukcji Faradaya i regułą Lenza. Ponieważ L_{21} jest stałe przy stałych cechach geometrycznych obu cewek i określonym ich położeniu względem siebie w ośrodku nieferromagnetycznym, więc korzystając ze wzoru (22.7) otrzymujemy

$$E_2 = -L_{21}\frac{dI_1}{dt}. \qquad (22.8)$$

Innymi słowy, *SEM indukcji wzajemnej* powstająca w cewce drugiej jest proporcjonalna do szybkości zmiany prądu w cewce pierwszej. SEM indukcji wzajemnej ma kierunek taki, że pole magnetyczne towarzyszące prądowi indukowanemu przeciwstawia się tym zmianom strumienia magnetycznego zewnętrznego, które wywołały powstanie prądu indukowanego.

Na podstawie wzoru (22.8) można podać drugą definicję indukcyjności wzajemnej. *Indukcyjność wzajemna* wyraża liczbowo SEM indukowaną w obwodzie drugim w wyniku jednostkowej zmiany natężenia prądu, zachodzącej w jednostce czasu w obwodzie pierwszym.

Jednostką indukcyjności wzajemnej w układzie SI *jest henr* (H). Indukcyjność wzajemna dwóch obwodów równa się 1 henrowi, gdy przy szybkości zmiany natężenia prądu w jednym obwodzie wynoszącej 1 A/s powstaje SEM indukcji wzajemnej równa 1 woltowi w obwodzie drugim

$$1\,\mathrm{H} = \frac{1\,\mathrm{V}}{1\,\mathrm{A/s}} = 1\frac{\mathrm{V}\cdot\mathrm{s}}{\mathrm{A}} = 1\,\Omega\cdot\mathrm{s}.$$

Henr jest jednostką stosunkowo dużą. W użyciu są jego podwielokrotne, jak milihenr i mikrohenr.

Rozważane zagadnienie można odwrócić: można szukać SEM indukowanej, powstającej w obwodzie *I* w wyniku zmian prądu w obwodzie *II*. Obowiązuje w tym przypadku analogiczny wzór:

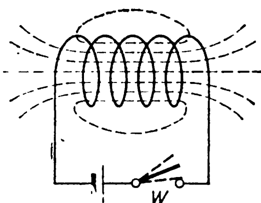
$$E_1 = -L_{12}\frac{dI_2}{dt}.$$

Oczywiście, w przypadku gdy w badanym obwodzie występuje SEM źródła prądu E_0 i SEM indukowana E_{ind}, obowiązuje zależność

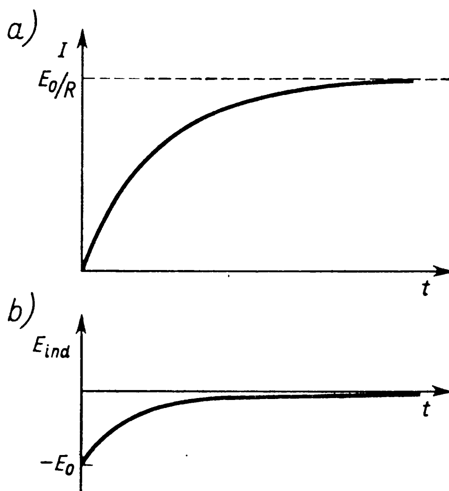
$$E_0 + E_{ind} = IR. \qquad (22.9)$$

22.5. Indukcja własna

Weźmy pod uwagę odosobniony obwód elektryczny, zawierający źródło prądu stałego, cewkę i wyłącznik *W* (rys. 22.11). Każda zmiana prądu w danym obwodzie powoduje w nim zmianę strumienia magnetycznego. Szybkość zmian strumienia magnetycznego decyduje — zgodnie z prawem Faradaya — o wartości liczbowej powstającej w danym obwodzie *SEM indukcji własnej*, zwanej też *SEM samoindukcji*. Kierunek tej SEM również jest zgodny z regułą Lenza.



Rys. 22.11



Rys. 22.12

Przeanalizujemy efekty indukcji własnej w rozważanym obwodzie w czasie: a) zamykania obwodu, b) przepływu prądu stałego, c) otwierania obwodu.

W przypadku a) zamknięcie obwodu wywołuje powstanie prądu płynącego z dowolnego źródła prądu stałego. Nazywać go będziemy dalej prądem bateryjnym. Powstaniu prądu bateryjnego towarzyszy pojawienie się strumienia magnetycznego i równocześnie — zgodnie z prawem Faradaya — pojawienie się w obwodzie prądu indukowanego (samoindukcyjnego) o kierunku przeciwnym niż prąd bateryjny. Wskutek sumowania się obu tych prądów prąd wypadkowy w obwodzie nie osiąga bezpośrednio po zamknięciu wyłącznika *W* wartości $I = E_0/R$ (gdzie E_0 — SEM źródła, R — opór

obwodu), wynikającej z prawa Ohma, lecz dopiero po pewnym czasie. Widać to z rys. 22.12a, przedstawiającego przebieg zależności prądu od czasu bezpośrednio po zamknięciu obwodu. Na rys. 22.12b przedstawione są równoczesne zmiany powstającej w tym obwodzie SEM indukcji własnej.

Od chwili osiągnięcia wartości $I = E_0/R$ prąd w obwodzie płynie jako prąd stały ($I =$ const, co odpowiada już przypadkowi b). Przepływowi takiego prądu towarzyszy stały strumień magnetyczny przechodzący przez dany obwód, nie istnieje zatem SEM indukcji własnej.

W chwili otwierania obwodu (przypadek c) prąd bateryjny powinien natychmiast spaść do zera. Szybkość zmiany strumienia magnetycznego jest tym razem bardzo duża. Powstaje zatem duża SEM indukcji własnej przeciwstawiająca się zaistniałym zmianom, a więc wywołująca prąd samoindukcyjny o kierunku zgodnym z prądem bateryjnym. Duża SEM indukcji własnej, pojawiająca się w chwili otwierania obwodów, powoduje często powstawanie iskry elektrycznej lub łuku elektrycznego między stykami wyłącznika W.

Jak widać z tych rozważań, oddziaływanie indukcyjne cewki w danym obwodzie na przebieg zjawisk elektrycznych w nim zachodzących nadaje prądowi elektrycznemu w obwodzie cechy bezwładności. Nawet w takim przypadku, gdy w obwodzie nie ma cewki (wielozwojowej), można go traktować jako cewkę o jednym zwoju. Zjawiska opisane wyżej zachodzą i w takim, jednozwojowym, obwodzie, przebieg ich jednak jest znacznie słabiej zaakcentowany.

Podsumowując wyniki rozważań powiemy, że dzięki zjawisku indukcji własnej obwód przeciwstawia się wszelkim zachodzącym w nim zmianom prądu. Po zamknięciu wyłącznika prąd w obwodzie narasta stopniowo od zera do wartości wynikającej z prawa Ohma, po otworzeniu wyłącznika — nie spada nieskończenie szybko do wartości zerowej.

Strumień magnetyczny przechodzący przez dany obwód jest w ośrodkach o stałej przenikalności magnetycznej proporcjonalny do prądu, a zatem

$$\Phi = LI.$$

Siła elektromotoryczna indukcji własnej równa się więc

$$E_{\text{ind wł}} = -\frac{d\Phi}{dt} = -L\frac{dI}{dt}, \tag{22.10}$$

gdzie L oznacza *indukcyjność własną obwodu* (zwaną także *współczynnikiem indukcji własnej*, *współczynnikiem samoindukcji*, lub najkrócej — *indukcyjnością obwodu*). Jednostką indukcyjności własnej w układzie SI jest znany już nam *henr*:

$$1\,\text{H} = \frac{1\,\text{V}}{1\,\text{A/s}} = 1\,\Omega\cdot\text{s}.$$

Indukcyjność własną 1 *henra* ma taki obwód, w którym jednostajna szybkość zmiany prądu równa 1 A/s wywołuje SEM samoindukcji równą 1 V.

Indukcyjność własna cewki zależy od liczby zwojów, gęstości ich nawinięcia, rozmiarów geometrycznych i kształtu cewki. Tak np. indukcyjność własna cewki (sole-

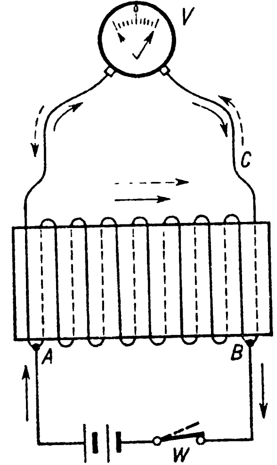
noidu) o długości l, powierzchni przekroju S i liczbie zwojów n na jednostkę długości równa się

$$L = \mu_0 n^2 l S,$$

gdzie μ_0 oznacza przenikalność magnetyczną próżni (w przybliżeniu i powietrza).

Dla cewek umieszczonych nie w powietrzu (próżni), lecz w dowolnym ośrodku o stałej przenikalności magnetycznej, w ostatnim wzorze występuje iloczyn $\mu_0 \mu_r$, gdzie μ_r jest ***przenikalnością magnetyczną względną*** danego ośrodka.

Objawy samoindukcji w obwodzie można wykazać w następującym doświadczeniu. Prąd z baterii (rys. 22.13) rozdziela się w punkcie A na dwie gałęzie: część płynie przez cewkę, reszta przez odnogę z woltomierzem. Kierunek prądu płynącego po zamknięciu wyłącznika W zaznaczony jest na rysunku strzałkami ciągłymi. Wskazówka miernika jest wychylona np. na prawo od zera w środku skali. W chwili otwarcia wyłącznika na końcach cewki powstaje SEM samoindukcji, podtrzymująca w zwojnicy prąd o kierunku zgodnym z pierwotnym, lecz płynący dalej przez punkt C i gałąź z miernikiem w kierunku przeciwnym do poprzedniego. Wskazówka miernika wychyla się przy tym w przeciwnym kierunku niż na początku.



Rys. 22.13

Przejdźmy obecnie do analizy matematycznej opisanych zjawisk.

W zamkniętym obwodzie zawierającym źródło prądu o SEM równej E_0 płynie prąd spełniający zależność:

$$E_0 + E_{\text{ind wł}} = RI,$$

$$E_0 - L\frac{dI}{dt} = RI.$$

Równanie to można sprowadzić do postaci

$$\frac{dI}{dt} + \frac{R}{L}\left(I - \frac{E_0}{R}\right) = 0.$$

Oznaczmy

$$I - \frac{E_0}{R} = x.$$

Wobec stałości E_0/R

$$dI = dx,$$

czyli

$$\frac{dx}{dt} + \frac{R}{L}x = 0,$$

$$\frac{dx}{x} = -\frac{R}{L}dt.$$

Po scałkowaniu otrzymujemy

$$\ln x = -\frac{R}{L}t + \text{const},$$

skąd

$$x = e^{-(R/L)t+\text{const}} = e^{-(R/L)t}e^{\text{const}} = Ce^{-(R/L)t}.$$

Po podstawieniu wartości x otrzymujemy

$$I = \frac{E_0}{R} + Ce^{-(R/L)t}. \qquad (22.11)$$

Wartość stałej całkowania C jest różna dla różnych warunków początkowych, a wraz z nią zmienia się przebieg zależności I od czasu. Sprawdzimy to dla dwóch przypadków.

1. Zamykanie obwodu; tym razem w chwili $t = 0$ prąd $I_0 = 0$. Z równania (22.11) wynika, że w tych warunkach

$$C = -\frac{E_0}{R},$$

a wzór na prąd przyjmuje postać

$$I = \frac{E_0}{R}(1 - e^{-(R/L)t}).$$

Innymi słowy, po zamknięciu obwodu prąd (zwany *prądem zamknięcia*) zmierza do wartości E_0/R, osiągając tę wartość teoretycznie po czasie $t = \infty$. Praktycznie zmiany I przedstawiają się następująco:

po czasie $t = \frac{L}{R}$ $\quad I = \frac{E_0}{R}\left(1 - \frac{1}{e}\right)$,

po czasie $t = \frac{2L}{R}$ $\quad I = \frac{E_0}{R}\left(1 - \frac{1}{e^2}\right)$ itd.

Po czasie $t = 5L/R$ prąd I różni się od E_0/R mniej niż o 1%. Wielkość L/R charakteryzuje szybkość narastania prądu, czyli stromość początkowej części krzywej z rys. 12.12a i nosi nazwę *stałej czasowej obwodu*. Im stosunek L/R jest mniejszy, tym szybciej osiąga się wartość $I \approx E_0/R$.

2. Rozważmy teraz przypadek otwierania obwodu, w którym płynie prąd $I = E_0/R$. W chwili $t = 0$ mamy zatem $I^{(0)} = I_0 = E_0/R$ i $E_0^{(0)} = 0$. W tych warunkach stała całkowania $C = I_0$ (por. (22.11)), a prąd wyraża się wzorem

$$I = I_0 e^{-(R/L)t} = \frac{E_0}{R} e^{-(R/L)t}. \qquad (22.12)$$

Tym razem prąd (zwany *prądem otwarcia*) maleje wykładniczo, przy czym znów szybkość spadku krzywej wykładniczej zależy od stałej czasowej obwodu L/R. Im stała czasowa jest mniejsza, tym krócej trwa spadek natężenia do zera przy otwieraniu ob-

wodu. Ogólnie biorąc, *im opór obwodu jest większy, a indukcyjność mniejsza, tym mniejszy jest wpływ samoindukcji na przebieg natężenia prądu.*

P y t a n i e. Stosunek L/R nazwaliśmy stałą czasową obwodu. Czy rzeczywiście wielkość tę mierzy się w jednostkach czasu?

22.6. Prąd sinusoidalnie zmienny

Omawiając efekty związane ze zjawiskiem indukcji elektromagnetycznej zetknęliśmy się z prądami zmieniającymi swe natężenie, jak również ze zmiennymi w czasie siłami elektromotorycznymi. Charakter tych zależności może być bardzo różny. Rozpatrzymy przypadek zależności sinusoidalnej, charakterystycznej dla prądnic prądu zmiennego, stosowanych w elektroenergetyce.

Siłę elektromotoryczną E prądu sinusoidalnie zmiennego jako funkcję czasu t można wyrazić zależnością

$$E = E_0 \sin\omega t, \tag{22.13}$$

gdzie E_0 oznacza *amplitudę siły elektromotorycznej*, zwaną też *napięciem szczytowym*, ω — pulsację zmian napięcia, ωt — fazę napięcia.

Warto zwrócić uwagę, że wypisana zależność formalnie zupełnie odpowiada zależności znanej nam już z teorii ruchu harmonicznego.

Najczęściej sinusoidalnie zmiennej sile elektromotorycznej E odpowiada sinusoidalnie zmienny prąd, który można wyrazić wzorem

$$I = I_0 \sin(\omega t + \varphi), \tag{22.14}$$

gdzie I_0 oznacza *amplitudę prądu, zwaną też prądem szczytowym*, ω — pulsację prądu, równą pulsacji napięcia, φ — przesunięcie w fazie przebiegu prądu względem przebiegu napięcia, a równocześnie fazę początkową prądu.

Jak widać z podanych wzorów, przebiegi czasowe I i E w prądach sinusoidalnie zmiennych można charakteryzować za pomocą amplitud prądu lub SEM oraz pulsacji ω (lub związanych z tą ostatnią częstotliwości ν lub okresu T; $\omega = 2\pi\nu = 2\pi/T$).

Okazało się jednak, że bardzo często zarówno przy obliczeniach obwodów elektrycznych prądu zmiennego, jak i w badaniach eksperymentalnych wygodnie jest posługiwać się innymi parametrami, a mianowicie tzw. wartościami skutecznymi prądu czy napięcia.

Wartość skuteczną danej wielkości (a więc w rozważanym przypadku wartość skuteczną prądu lub SEM) definiujemy jako taką stałą wartość rozważanej wielkości (I_{sk} lub E_{sk}), która w obwodzie o oporze R w ciągu jednego okresu T lub jego wielokrotności powoduje wydzielenie tej samej ilości ciepła (czyli ten sam *skutek* cieplny) jak badana wielkość zmienna.

Stosując tę definicję np. do obwodu prądu sinusoidalnie zmiennego w odniesieniu do jednego okresu T znajdziemy I_{sk} z następującej zależności:

$$I_{sk}^2 RT = \int_0^T RI_0^2 \sin^2(\omega t + \varphi)\, dt.$$

Korzystając z zależności

$$\sin^2(\omega t+\varphi) = \frac{1-\cos 2(\omega t+\varphi)}{2}$$

otrzymamy

$$\int_0^T \sin^2(\omega t+\varphi)\,dt = \tfrac{1}{2}\int_0^T dt - \tfrac{1}{2}\int_0^T \cos 2(\omega t+\varphi)\,dt = \frac{T}{2},$$

a zatem

$$I_{sk}^2 T = I_0^2 \frac{T}{2},$$

$$I_{sk} = I_0/\sqrt{2} \approx 0{,}7 I_0.$$

Rozumując podobnie w odniesieniu do E_{sk} znajdziemy

$$\frac{E_{sk}^2}{R} T = \frac{E_0^2}{R}\int_0^T \sin^2\omega t\,dt = \frac{E_0^2}{R}\frac{T}{2},$$

$$E_{sk} = E_0/\sqrt{2} \approx 0{,}7 E_0.$$

Dane liczbowe charakteryzujące prąd zmienny najczęściej dotyczą właśnie wartości skutecznych. Tak np. mówiąc, że napięcie prądu zmiennego w sieci miejskiej wynosi 220 V, mamy na myśli wartość skuteczną E_{sk}. Amplituda SEM w sieci miejskiej wynosi zatem $E_0 \approx 310$ V.

Interesujące jest zagadnienie *średniej mocy prądu sinusoidalnie zmiennego*. Średnia moc P, *odniesiona do jednego okresu* T, wyraża się wzorem

$$P = \frac{\int_0^T E_0 I_0 \sin\omega t \sin(\omega t+\varphi)\,dt}{T} = \frac{E_0 I_0}{T}\int_0^T (\sin^2\omega t\cos\varphi + \sin\omega t\cos\omega t\sin\varphi)\,dt,$$

skąd

$$P = \tfrac{1}{2} E_0 I_0 \cos\varphi$$

lub

$$P = E_{sk} I_{sk} \cos\varphi.$$

Wzór ten wyraża tzw. *moc czynną prądu zmiennego*. Jak widać, jest ona zależna od różnicy faz między przebiegiem napięcia i prądu. Zmiana kąta φ przy tych samych wartościach I_{sk} i E_{sk} pociąga za sobą zmianę mocy czynnej. Moc czynna jest największa, gdy nie ma różnicy faz między przebiegiem prądu i napięcia, natomiast równa się zeru, gdy $\varphi = \pi/2$. W tym ostatnim przypadku mówimy o *prądzie biernym lub prądzie bezwatowym*. W następnym paragrafie zobaczymy, jakie czynniki w obwodzie prądu zmiennego wpływają na różnicę faz φ.

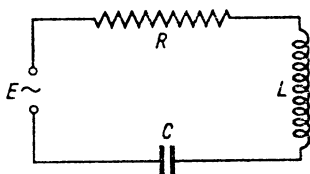
22.7. Natężenie prądu zmiennego jako funkcja indukcyjności, pojemności i oporu

Szerokie omówienie prądów sinusoidalnie zmiennych przewidziane jest w kursie podstaw elektrotechniki, toteż tutaj ograniczymy się tylko do kilku uwag wskazujących na pewne różnice w zjawiskach zachodzących w obwodach prądu sinusoidalnie zmien-

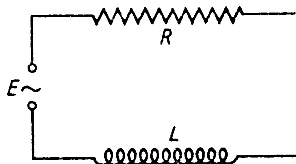
nego i prądu o stałym kierunku i o stałej wartości I. Podamy też zależności matematyczne, obowiązujące w odniesieniu do prądów zmiennych.

W przypadku prądu stałego ($I = \text{const}$) włączenie do obwodu elektrycznego pojemności C (kondensatora) wyklucza możliwość płynięcia w nim prądu o stałej wartości I. Indukcyjność L obwodu w tym przypadku nie odgrywa roli, gdyż $dI/dt = 0$.

W przypadku prądu sinusoidalnie zmiennego włączenie pojemności nie uniemożliwia przepływu prądu, ale ma wpływ na wartość I i na różnicę faz między przebiegami prądu i napięcia. Ponadto istotną rolę odgrywa indukcyjność L obwodu, gdyż każdej zmianie prądu towarzyszy zmiana strumienia magnetycznego, a więc w obwodzie prądu zmiennego, zgodnie z regułą Lenza, powstaje dodatkowa siła elektromotoryczna SEM o kierunku zmiennym. Przebieg prądu zmiennego jest zatem funkcją R, L i C. Ograniczymy się do podania I jako funkcji R, L, C dla kilku typowych obwodów prądu zmiennego.



Rys. 22.14



Rys. 22.15

Przypadek najogólniejszy: do obwodu włączony jest opór R, zwany ***rezystancją***, pojemność C i samoindukcyjność L (rys. 22.14). Jest to zatem *obwód RLC*. Włączona SEM zmienia się według (22.13). W tym przypadku

$$I = \frac{E_0}{\sqrt{R^2 + \left(L\omega - \frac{1}{C\omega}\right)^2}} \sin(\omega t - \varphi). \qquad (22.15)$$

Opór Z obwodu prądu elektrycznego zmiennego nosi nazwę ***impedancji*** i wyraża się następującą liczbą zespoloną:

$$Z = R + i\left(L\omega - \frac{1}{C\omega}\right).$$

Moduł tej wielkości zespolonej, czyli tzw. moduł impedancji, ma postać

$$\sqrt{R^2 + \left(L\omega - \frac{1}{C\omega}\right)^2}.$$

Jest to właśnie wielkość występująca w (22.15), decydująca o natężeniu prądu zmiennego w obwodzie, do którego włączona jest SEM wyrażona wzorem (22.13). *Impedancja i jej moduł zależą od rezystancji R, reaktancji indukcyjnej $L\omega$ i reaktancji pojemnościowej $1/C\omega$.*

W rozważanym wyżej przypadku tangens różnicy faz przebiegu prądu i napięcia wyraża się wzorem

$$\operatorname{tg}\varphi = \frac{L\omega - 1/C\omega}{R}. \tag{22.16}$$

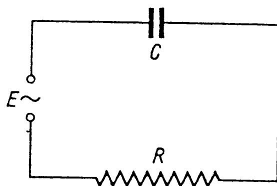
Różnica faz φ jest zatem funkcją L, C, R i pulsacji ω. Może ona przyjmować wartości zarówno dodatnie, jak i ujemne, czyli przebieg prądu może wyprzedzać fazę lub spóźniać się w fazie względem przebiegu napięcia zależnie od wartości L i C.

Z podanego wzoru wynikają zależności obowiązujące w przypadkach szczególnych. Niech obwód zawiera rezystancję R i indukcyjność L (rys. 22.15). Jest to zatem *obwód RL*. Do obwodu włączona jest, jak poprzednio, SEM zmieniająca się według równania (22.13).

W tym przypadku wzór (22.15) redukuje się do postaci

$$I = \frac{E_0}{\sqrt{L^2\omega^2 + R^2}} \sin(\omega t - \varphi). \tag{22.17}$$

Czynnik $\sqrt{L^2\omega^2+R^2}$ występujący w mianowniku przedstawia moduł oporu obwodu RL prądu sinusoidalnie zmiennego zależny, jak widać, od rezystancji R i reaktancji indukcyjnej $L\omega$ (pojemność $C = \infty$).



Rys. 22.16

Różnica faz przebiegów prądu i napięcia spełnia zależność

$$\operatorname{tg}\varphi = \frac{L\omega}{R}, \tag{22.18}$$

czyli $\operatorname{tg}\varphi$ równa się stosunkowi reaktancji indukcyjnej do rezystancji. W tym przypadku przebieg prądu spóźnia się w fazie względem przebiegu napięcia.

Innym przypadkiem szczególnym będzie przepływ prądu sinusoidalnie zmiennego w obwodzie zawierającym pojemność C i rezystancję R (tzw. *obwód RC*). Siła elektromotoryczna zmienia się jak poprzednio. Tym razem (rys. 22.16):

$$I = \frac{E_0}{\sqrt{R^2 + \left(\frac{1}{C\omega}\right)^2}} \sin(\omega t + \varphi), \tag{22.19}$$

przy czym

$$\operatorname{tg}\varphi = \frac{1}{C\omega R}.$$

Obecnie natężenie prądu wyprzedza w fazie siłę elektromotoryczną. Różnica faz φ jest funkcją pojemności C, rezystancji R i pulsacji ω. tg φ równa się stosunkowi reaktancji pojemnościowej do rezystancji.

Analizując bliżej wzór na prąd I w obwodzie RL widzimy, że impedancja rośnie ze wzrostem pulsacji ω. Innymi słowy, przy dużych częstotliwościach — nawet w przypadku małej rezystancji R — otrzymujemy małe prądy. Na tym polega działanie tzw. dławików, tzn. cewek o dużej indukcyjności, które hamują przepływ prądu o dużej częstotliwości. Dławik jest zatem filtrem przepuszczającym tylko prądy o małych częstotliwościach.

Inaczej jest w obwodach RC, gdy w obwodzie mamy rezystancję i reaktancję pojemnościową. Ze wzrostem ω przy danym C impedancja, równa tym razem $\sqrt{R^2+(1/C\omega)^2}$, maleje. W obwodzie powstają prądy dużej częstotliwości o znacznym natężeniu. Odwrotnie — w przypadku prądów o małej częstotliwości impedancja jest duża i prądy niskiej częstotliwości mają małe natężenie.

Pytania i zadania

1. W jakich warunkach z określoną powierzchnią S związany jest a) stały, b) zmienny w czasie strumień magnetyczny?

2. Na czym polega i w jakich warunkach powstaje zjawisko indukcji elektromagnetycznej?

3. W jakich obwodach elektrycznych powstają prądy indukowane, a w jakich pojawiają się jedynie SEM indukowane?

4. Co decyduje o wartości powstającej SEM indukowanej i o wartości prądu indukowanego?

5. Sformułować regułę Lenza. Wyprowadzić prawo indukcji elektromagnetycznej Faradaya.

6. Jaka indukowana siła elektromotoryczna powstaje w przewodniku o długości l, poruszającym się z prędkością $\mathbf{v}$ w jednorodnym polu magnetycznym o indukcji $\mathbf{B}$ prostopadłej do $\mathbf{v}$?

7. Scharakteryzować prądy wirowe Foucaulta. Jakie są sposoby ich zmniejszania? Jak można je wykorzystywać?

8. Omówić warunki występowania zjawiska indukcji wzajemnej. Od czego zależy siła elektromotoryczna indukcji wzajemnej? Jaki jest kierunek prądu indukcyjnego, powstającego w wyniku indukcji wzajemnej dwóch obwodów?

9. Jaki jest sens fizyczny współczynnika indukcji wzajemnej dwóch obwodów. Co jest jednostką indukcyjności? Podać definicję tej jednostki.

10. Przeanalizować zjawisko indukcji własnej. Podać wykresy zależności natężenia prądu i siły elektromotorycznej indukcji własnej od czasu, wyjaśniając szczegółowo warunki związane z każdą częścią wykresu.

11. Uzasadnić twierdzenie, że samoindukcja nadaje prądowi w obwodzie elektrycznym cechę bezwładności.

12. Jakiemu obwodowi elektrycznemu przypisuje się indukcyjność równą 1 H?

13. Jak się przedstawia analiza matematyczna zjawiska samoindukcji? Jakie jest matematyczne uzasadnienie faktu, że dla różnych założonych warunków początkowych otrzymujemy różne wyrażenia na zależność prądu od czasu?

14. Jaka jest charakterystyka a) prądu zamknięcia, b) prądu otwarcia.

15. Co to jest stała czasowa obwodu elektrycznego?

16. Jaki prąd nazywamy prądem sinusoidalnie zmiennym? Jakie wzory wyrażają w tym przypadku natężenie prądu I i SEM w zależności od czasu?

17. Co nazywamy szczytowymi wartościami prądu i napięcia?

18. Rozróżnić wartości szczytowe i skuteczne prądu zmiennego. Podać zależności matematyczne wiążące obie wielkości odniesione do prądu i napięcia.

19. Czym się mogą różnić obwody prądu stałego (I = const) i zmiennego? Co nazywamy impedancją? Jakie składniki się na nią składają?

20. Czy reaktancja indukcyjna i pojemnościowa zależą od częstotliwości zmian prądu zmiennego? Jeśli tak, to omówić odpowiednie zależności.

21. Jakim wzorem wyraża się moc prądu sinusoidalnie zmiennego? Na czym polega różnica w porównaniu z mocą prądu stałego?

22. Indukcja magnetyczna jednorodnego pola magnetycznego równa się 20 T. Jaki strumień magnetyczny przypada na powierzchnię 1 dm^2, jeżeli normalna do tej powierzchni tworzy kąt 30° z kierunkiem **B**?

Odp. 0,17 Wb.

23. Przewodnik o długości $l = 60$ cm porusza się prostopadle do kierunku linii jednorodnego pola magnetycznego o indukcji $B = 0{,}55$ T z prędkością $v = 2$ m/s. Jaka SEM indukuje się w tym przewodniku?

Odp. 0,66 V.

24. Przewodnik składa się z 3 kołowych zwojów o promieniu $r = 10$ cm, umieszczonych prostopadle do linii jednorodnego pola magnetycznego o indukcji $B = 1\ Wb/m^2$. Jaka SEM indukuje się w tym przewodniku, jeśli pole magnetyczne zanika jednostajnie w ciągu 0,5 s.

Odp. 0,189 V.

25. Jaka średnia SEM indukuje się przy półobrocie ramki prostokątnej o bokach $a = 30$ cm i $b = 20$ cm obracającej się dokoła boku a prostopadłego do linii pola o natężeniu $H = \frac{10^3}{4\pi}\ 5000\ A \cdot m^{-1}$ z częstotliwością $\nu = 60\ s^{-1}$?

Odp. 7,2 V.

26. Jaka SEM indukuje się w prostoliniowym odcinku przewodnika o długości $l = 30$ cm, jeśli przewodnik ten obraca się z częstotliwością $\nu = 30\ s^{-1}$ w jednorodnym polu magnetycznym o indukcji $B = 0{,}5\ Wb/m^2$ w płaszczyźnie prostopadłej do linii pola, dokoła osi przechodzącej przez koniec przewodnika?

Odp. 4,24 V.

27. W solenoidzie mającym średnicę $2r = 10$ cm i liczącym $n = 1000$ zwojów na swej długości $l = 50$ cm natężenie prądu równomiernie wzrasta o 0,1 A na każdą sekundę. Na ten solenoid nasunięto pojedynczy zwój z przewodnika miedzianego o przekroju $S = 2\ mm^2$. Zakładając, że strumienie indukcji przenikające przez solenoid i przez przewód miedziany są w każdej chwili jednakowe, obliczyć natężenie prądu indukowanego w przewodniku miedzianym? Opór właściwy miedzi $\varrho = 1{,}7\ \Omega \cdot m$.

Odp. $7{,}4 \cdot 10^{-4}$ A.

28. Jaka jest SEM w cewce o indukcyjności $L = 0{,}05$ H, gdy natężenie prądu wzrasta w każdej sekundzie o 10 A?

Odp. 0,5 V.

29. Cewka złożona z 10 000 zwojów o średnicy 7 cm i długości 50 cm zrobiona jest z miedzianego przewodnika o przekroju 0,1 mm^2. Po jakim czasie od chwili zamknięcia wyłącznika prąd w obwodzie cewki osiągnie wartość równą 0,99 wartości maksymalnej?

Odp. 0,0119 s.

30. W cewce, której rezystancja równa się 10 Ω, obserwuje się przesunięcie fazowe prądu i napięcia równe 60°. Jaka jest indukcyjność cewki, jeżeli częstotliwość prądu równa się 50 Hz?

Odp. 0,054 H.

31. Jakie jest maksymalne natężenie prądu w obwodzie złożonym z pojemności $C = 20\ \mu F$ i rezystancji $R = 150\ \Omega$, jeżeli zastosowano napięcie $E_0 = 220$ V prądu sinusoidalnie zmiennego o częstotliwości 50 Hz?

Odp. Około 1 A.

ROZDZIAŁ 23

Drgania i fale elektromagnetyczne

W rozdziale niniejszym zajmować się będziemy procesami, w których następują kolejne przemiany energii E_e pola elektrycznego w energię E_m pola magnetycznego i odwrotnie. Z przemianami takimi zapoznamy się rozważając drgania elektromagnetyczne, zachodzące w obwodach elektrycznych *LC* i *LRC* (tzn. złożonych z indukcyjności i pojemności lub z indukcyjności, rezystancji i pojemności) oraz rozważając promieniowanie drgającego dipola elektrycznego rozchodzące się w przestrzeni w postaci fali elektromagnetycznej.

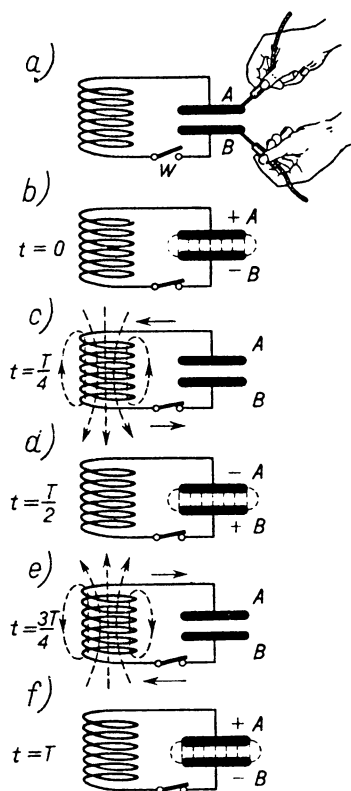
23.1. Analiza jakościowa drgań elektromagnetycznych

Analizę jakościową drgań elektromagnetycznych przeprowadzimy na przykładzie drgań powstających podczas rozładowania kondensatora o pojemności C w obwodzie zawierającym cewkę o indukcyjności L (rys. 23.1a). Przez połączenie okładek A i B kondensatora z biegunami źródła prądu stałego (przy otwartym wyłączniku W) ładujemy okładkę A dodatnio, B — ujemnie. Wskutek tego między okładkami kondensatora wytwarza się pole elektryczne, które — jak pamiętamy — związane jest z występowaniem energii elektrycznej E_e (rys. 23.1b).

Po odłączeniu kondensatora od źródła prądu stałego i zamknięciu obwodu, kosztem energii E_e zaczyna płynąć prąd od okładki A przez cewkę do okładki B zwiększając swe natężenie stopniowo dzięki istnieniu w obwodzie indukcyjności L. Przepływowi prądu elektrycznego towarzyszy pojawienie się pola magnetycznego, które związane jest z występowaniem energii magnetycznej E_m. Pole to jest zlokalizowane w zasadzie w obszarze cewki. Przepływ ładunków od jednej okładki do drugiej doprowadza w pewnej chwili do spadku napięcia między okładkami do zera, a więc do zaniku pola elektrycznego i związanej z nim energii (rys. 23.1c). W tej samej chwili energia E_m osiąga wartość maksymalną.

W dalszym etapie zjawiska, dzięki SEM indukcji własnej w cewce prąd płynie nadal w tym samym kierunku zmniejszając swe natężenie do zera. Równocześnie okładka B ładuje się dodatnio, okładka A — ujemnie. Znów pojawia się pole elektryczne, a wraz z nim energia E_e powstająca w tej części procesu kosztem energii E_m. Rys. 23.1d

odpowiada chwili, gdy w obwodzie nie ma już prądu elektrycznego (a więc energia magnetyczna równa się zeru), lecz istnieje pole elektryczne o maksymalnym natężeniu (energia elektryczna $E_e = E_{e\,max}$) i kierunku przeciwnym niż na początku procesu przedstawionym na rys. 23.1b.



Rys. 23.1

W dalszym przebiegu zjawiska obserwuje się przepływ prądu od okładki B do A, przy czym kosztem energii E_e pojawia się energia E_m osiągająca maksimum wtedy, gdy pole elektryczne a z nim razem i energia E_e spadają do zera (rys. 23.1e). Następnie znów dzięki SEM indukcji własnej osiągnięty zostaje stan zerowej energii E_m (brak prądu) i maksymalnej energii E_e (rys. 23.1f), czyli stan równoważny stanowi początkowemu, przedstawionemu na rys. 23.1b.

Całość omówionego wyżej zjawiska może powtarzać się wielokrotnie, podobnie jak wielokrotnie odbywają się drgania cieczy doskonałej dokoła położenia równowagi w naczyniu połączonym lub drgania wahadła w próżni. Wykorzystując tę analogię mówimy o drganiach elektromagnetycznych w obwodzie. Dla pogłębienia analogii warto zwrócić uwagę na stronę energetyczną obu zjawisk. W drganiach mechanicznych harmonicznych, którym poświęciliśmy § 10.1, występowały przemiany energetyczne: kosztem energii mechanicznej potencjalnej punktu drgającego wytwarzała się energia kinetyczna i odwrotnie. Gdy jedna z nich osiągała maksimum, druga miała wartość zerową. Spoglądając na rys. 23.1b–f stwierdzamy, że w rozważanym cyklu przemian (w ciągu okresu T) w obwodzie elektrycznym mamy zmiany energii E_e i E_m o podobnym przebiegu, a mianowicie:

t	0	$T/4$	$T/2$	$3T/4$	T
E_e	max	0	max	0	max
E_m	0	max	0	max	0

Do bliższego określenia charakteru zmian energii E_e i E_m, w zależności od czasu, wrócimy jeszcze przy końcu p. 23.2.2.

Jak zobaczymy dalej, analogia między drganiami mechanicznymi i elektromagnetycznymi jest jeszcze szersza. *W dziedzinie drgań elektromagnetycznych* można będzie — podobnie jak w dziedzinie drgań mechanicznych — wyróżnić *drgania nietłumione* w obwodach bez strat energii (np. w obwodzie modelowym LC), *drgania tłumione* (w obwodach LRC) oraz *drgania wymuszone* przy zastosowaniu wymuszającej periodycznie zmiennej siły elektromotorycznej.

23.2. Analiza ilościowa drgań elektromagnetycznych

Przeanalizujmy drgania elektromagnetyczne zachodzące podczas rozładowania kondensatora o pojemności C w układzie szeregowym, zawierającym indukcyjność L i rezystancję R. Analizę przeprowadzimy pod kątem zmienności prądu w obwodzie w zależności od czasu, aczkolwiek można by ją także prowadzić z punktu widzenia zmienności ładunku Q na okładce kondensatora lub napięcia U między okładkami.

Zgodnie z drugim prawem Kirchhoffa suma czynnych w obwodzie sił elektromotorycznych równa się sumie spadków napięcia, a zatem

$$-L\frac{dI}{dt} = U+RI, \tag{23.1}$$

gdzie wyraz pierwszy oznacza siłę elektromotoryczną indukowaną, drugi — spadek napięcia na kondensatorze, trzeci — spadek napięcia na rezystancji.

Celem wyrugowania wielkości U korzystamy z zależności (por. § 19.11):

$$U = Q/C = \int I dt/C. \tag{23.2}$$

Wtedy (23.1) przyjmuje postać

$$-L\frac{dI}{dt} = \frac{\int I dt}{C}+RI. \tag{23.3}$$

Po zróżniczkowaniu otrzymujemy

$$-L\frac{d^2I}{dt^2} = \frac{I}{C}+R\frac{dI}{dt}$$

lub

$$\frac{d^2I}{dt^2}+\frac{R}{L}\frac{dI}{dt} = -\frac{1}{LC}I. \tag{23.4}$$

W następnych punktach (23.2.1) i (23.2.2) rozpatrzymy oddzielnie dwa przypadki: 1) gdy rezystancja R obwodu jest tak mała, że można ją zaniedbać (układ LC), 2) gdy rezystancji R obwodu nie można zaniedbać (układ LRC).

23.2.1. *Drgania elektromagnetyczne w układzie LC.* Gdy rezystancja R badanego obwodu jest tak mała, że można ją zaniedbać, równanie (23.4) redukuje się do postaci

$$\frac{d^2I}{dt^2} = -\frac{1}{LC}I. \tag{23.5}$$

Równanie to jest analogiczne do równania różniczkowego ruchu harmonicznego nietłumionego (por. wzór 10.9):

$$\frac{d^2s}{dt^2} = -\omega_0^2 s \tag{23.6a}$$

(przyspieszenie d^2s/dt^2 jest proporcjonalne do wychylenia s od położenia równowagi; współczynnik proporcjonalności ω_0^2 — kwadrat pulsacji drgań własnych układu — równa się $4\pi^2/T^2$).

Rozwiązanie tego równania, czyli wyrażenie na wychylenie od położenia równowagi, ma postać

$$s = A\sin(\omega_0 t+\varphi), \tag{23.6b}$$

gdzie A oznacza amplitudę wychylenia, φ — fazę początkową. Analogiczne równanie wyraża prąd I jako funkcję czasu

$$I = I_0\sin(\omega_0 t+\varphi),$$

gdzie I_0 oznacza amplitudę prądu, ω_0 — jego pulsację, φ — fazę początkową.

Wartość fazy początkowej znajdujemy ustalając warunki początkowe rozpatrywanego zjawiska. W chwili $t = 0$ mieliśmy najsilniejsze naładowanie kondensatora, lecz w obwodzie nie było prądu, czyli $I_{t=0} = 0$. Stąd wynika, że $\varphi = 0$, a zatem natężenie prądu rozładowania kondensatora w układzie LC wyraża się wzorem:

$$I = I_0\sin\omega_0 t. \tag{23.7}$$

Przebieg prądu ma więc charakter sinusoidalny, zgodny z (23.7).

Z przyrównania wartości współczynników proporcjonalności w równaniach (23.5) i (23.6a) otrzymujemy

$$\omega_0^2 = \frac{1}{LC}, \tag{23.8}$$

skąd wynika zależność:

$$\frac{4\pi^2}{T^2} = \frac{1}{LC}.$$

Po przekształceniu tego wyrażenia otrzymujemy *wzór Thomsona* na okres drgań własnych (nietłumionych) elektromagnetycznych obwodu LC:

$$T = 2\pi\sqrt{LC}, \tag{23.9}$$

gdzie okres T wyrażony jest w sekundach, indukcyjność L w henrach, pojemność C w faradach.

Celem znalezienia amplitudy I_0 rozważanych drgań podstawiamy do wzoru (23.1)

$$\frac{dI}{dt} = I_0\omega_0\cos\omega_0 t$$

obliczone z (23.7). Uwzględniając założenie, że $R = 0$ i przyjmując jako warunek początkowy, że w chwili $t = 0$ napięcie na kondensatorze $U = U_0$, gdzie U_0 oznacza amplitudę napięcia na kondensatorze, znajdujemy

$$U_0 = LI_0\omega_0,$$

skąd

$$I_0 = \frac{U_0}{L\omega_0} = \frac{U_0}{\sqrt{L/C}}.$$

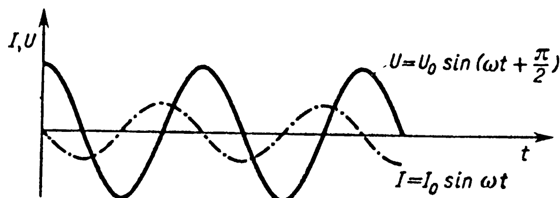
Uwzględniając zależność:

$$Q = CU$$

i wynikającą z niej po zróżniczkowaniu:

$$I = \frac{dQ}{dt} = C\frac{dU}{dt}$$

stwierdzamy proporcjonalność I do pochodnej napięcia. A zatem, sinusoidalnej zmienności prądu w zależności od czasu odpowiada cosinusoidalna zależność napięcia od czasu. Innymi słowy, gdy zaniedbujemy rezystancję obwodu, między prądem i napięciem istnieje przesunięcie fazowe równe $\pi/2$ (rys. 23.2).

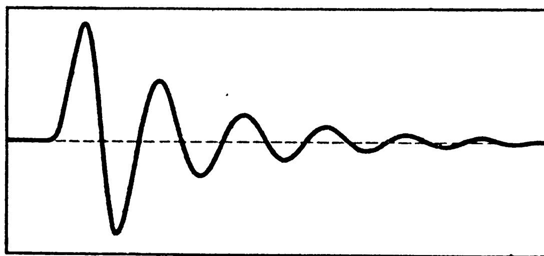


Rys. 23.2

Na tym kończymy analizę ilościową drgań elektromagnetycznych nietłumionych, odbywających się w obwodzie *LC*.

23.2.2. *Drgania elektromagnetyczne w obwodzie LRC*. Gdy rezystancji R obwodu nie można zaniedbać, rozwiązaniem równania (23.4) (analogicznego do równania różniczkowego drgań mechanicznych tłumionych) jest równanie drgań elektromagnetycznych tłumionych. W przypadku obwodów o małym tłumieniu przez analogię do wzoru (10.20) rozwiązanie to można napisać w postaci:

$$I = \frac{U_0}{\omega_0 L} e^{-\delta t} \sin\omega_1 t, \tag{23.10}$$



Rys. 23.3

gdzie δ oznacza współczynnik, zwany ***stałą tłumienia***, decydujący o szybkości zanikania drgań (założenie: δ — bardzo małe), ω_1 — pulsację drgań zanikających. Pozostałe symbole mają znaczenie takie samo, jak poprzednio. Wyraz $U_0/\omega_0 L$ oznacza maksymalną wartość prądu. W tym przypadku przebieg zmian prądu w drganiach elektromagnetycznych tłumionych w zależności od czasu przedstawia rys. 23.3.

Celem bliższego określenia wielkości δ i ω_1 skorzystamy raz jeszcze z analogii do drgań mechanicznych tłumionych. Stała tłumienia drgań mechanicznych spełnia zależność $\delta = b/2m$, gdzie b oznacza współczynnik oporu, m — masę, czyli miarę bezwładności. Stała tłumienia stanowi też połowę współczynnika przy pierwszej pochodnej wychylenia od położenia równowagi względem czasu we wzorze (10.19). W rozważanym obwodzie elektrycznym miarą oporu stawianego prądowi jest rezystancja R, a miarą „bezwładności" — jak to już wspominaliśmy (por. § 22.5) — jest indukcyjność L. A zatem stosując analogię powiemy, że w przypadku drgań elektromagnetycznych

$$\delta = R/2L. \tag{23.11}$$

Łatwo sprawdzić, że R/L stanowi też współczynnik przy dI/dt w równaniu (23.4).

Pulsację drgań mechanicznych tłumionych wyraża wzór:

$$\omega_1 = \sqrt{\omega_0^2 - \delta^2}.$$

Ten sam wzór odnosi się do drgań elektromagnetycznych tłumionych. Po podstawieniu ω_0 i δ otrzymujemy

$$\omega_1 = \sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}}. \tag{23.12}$$

Szybkość zanikania drgań elektromagnetycznych można również scharakteryzować za pomocą znanego już nam *logarytmicznego dekrementu tłumienia* Λ. Definicja tej wielkości jest w zasadzie taka sama jak w przypadku drgań mechanicznych z tą tylko różnicą, że jest to logarytm naturalny stosunku amplitud prądu elektrycznego obliczonych dla czasów t i $(t+T_1)$, gdzie T_1 jest okresem drgań tłumionych. A zatem

$$\Lambda = \ln \frac{e^{-\delta t}}{e^{-\delta(t+T_1)}} = \delta T_1.$$

Uwzględniając (23.11) i podstawiając zamiast okresu T_1 przybliżoną dla drgań tłumionych wartość okresu drgań własnych układu, wynikającą ze wzoru Thomsona (por. zad. 18), otrzymujemy

$$\Lambda \cong \pi \frac{R}{\sqrt{L/C}}. \tag{23.13}$$

Większym wartościom logarytmicznego dekrementu tłumienia (i stałej tłumienia) odpowiada szybszy zanik drgań.

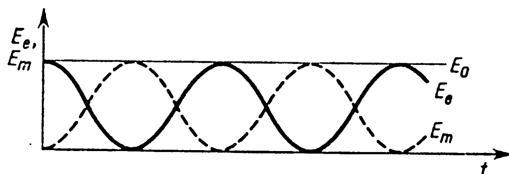
Z równania (23.13) widać, że szybkiemu zanikowi drgań elektromagnetycznych sprzyjają przede wszystkim duża rezystancja R oraz — aczkolwiek w mniejszym stopniu — duża pojemność C i mała indukcyjność L.

Biorąc pod uwagę (23.12) można określić warunek, który musi być spełniony, aby drgania tłumione miały charakter periodyczny. Zachodzi to wtedy, gdy

$$1/LC > R^2/4L^2, \quad \text{czyli gdy} \quad R < 2\sqrt{L/C}.$$

Rozważania nasze dotyczące drgań nietłumionych i tłumionych zakończymy ponownym akcentem energetycznym. Wiemy już z § 23.1, że w czasie rozładowywania

kondensatora w obwodzie LC ($R = 0$) zachodzą przemiany energii pola elektrycznego E_e na energię pola magnetycznego E_m. Wobec pełnej analogii do drgań harmonicznych nietłumionych zmiany obu energii w czasie mają ten sam charakter co i zmiany energii mechanicznej potencjalnej E_p i kinetycznej E_k. Krzywe przedstawiające zmiany obu energii w zależności od czasu podczas drgań elektromagnetycznych (rys. 23.4) są co do kształtu identyczne z krzywymi dotyczącymi energii drgań mechanicznych (rys. 10.4). Z obu rysunków widać, że energia całkowita E_0 w obu przypadkach jest stała.



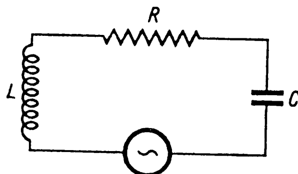
Rys. 23.4

Gdy rezystancji R (i wszelkich innych czynników powodujących tłumienie) nie można zaniedbać, energia całkowita maleje, podobnie jak to ma miejsce w układach mechanicznych z tłumieniem.

Kompensując straty energii przez dostarczanie jej w odpowiedniej ilości z zewnętrznego źródła można otrzymać drgania nie zanikające w układzie LRC. Odpowiednie urządzenie stanowi generator drgań elektrycznych o pulsacji stałej lub regulowanej — zależnie od parametrów części składowych.

23.3. Drgania elektromagnetyczne wymuszone. Rezonans

W rozdziale 10 nie poprzestaliśmy na omówieniu drgań harmonicznych własnych nietłumionych i tłumionych, lecz — przez uwzględnienie działania sinusoidalnie zmiennej siły $F = F_0 \sin \Omega t$ (por. § 10.5) — rozszerzyliśmy zakres rozważań na drgania wymuszone i ściśle z nim związane zjawisko rezonansu.



Rys. 23.5

Okazuje się, że w obwodach elektrycznych można również wywołać drgania wymuszone. Wystarczy np. w tym celu do szeregowego obwodu LRC (tzn. obwodu, w którym L, R i C połączone są szeregowo) włączyć źródło sinusoidalnie zmiennego napięcia (rys. 23.5). Wskutek tego nasze podstawowe równanie (23.1) rozszerzy się o dodatkowy wyraz odpowiadający sinusoidalnym zmianom SEM włączonego źródła:

$$E = E_0 \sin \Omega t$$

(gdzie E_0 oznacza amplitudę SEM, Ω — pulsację SEM), przyjmując postać:

$$-L\frac{dI}{dt}+E_0\sin\Omega t = U+RI. \tag{23.14}$$

Nie wnikając w szczegóły dalszych rozważań związanych z uzyskaniem nowego równania różniczkowego ze zmiennymi I i t oraz jego rozwiązaniem podkreślimy najważniejsze wyniki: po pewnym czasie od chwili zastosowania wymuszającej SEM ustalony prąd wymuszony w obwodzie a) ma przebieg sinusoidalny o pulsacji Ω równej pulsacji wymuszającej SEM, b) wykazuje przesunięcie fazowe względem SEM, c) ma amplitudę I_0 wyrażoną równaniem:

$$I_0 = \frac{E_0}{\sqrt{(\Omega L-1/\Omega C)^2+R^2}}. \tag{23.15}$$

Rozpatrzmy bliżej ostatnie wyrażenie. Jak widać, amplituda I_0 jest funkcją Ω. Prąd wymuszony o maksymalnej amplitudzie, czyli prąd odpowiadający zjawisku rezonansu, w szeregowym obwodzie LRC otrzymamy wtedy, gdy

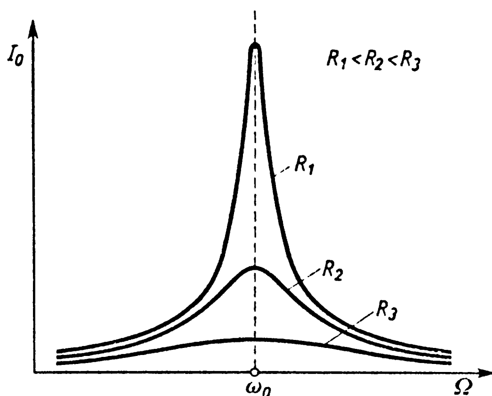
$$\Omega_r L-1/\Omega_r C = 0 \tag{23.16}$$

(gdzie Ω_r oznacza pulsację rezonansową wymuszającej SEM), czyli gdy

$$\Omega_r = \sqrt{1/LC}. \tag{23.17}$$

Uwzględnienie (23.8) sprowadza warunek rezonansu dla obwodów elektrycznych do postaci

$$\Omega_r = \omega_0^*. \tag{23.18}$$



Rys. 23.6

Innymi słowy, *prąd wymuszony osiąga maksymalną, rezonansową amplitudę, gdy pulsacja wymuszającej SEM jest równa pulsacji drgań własnych nietłumionych układu.*

---

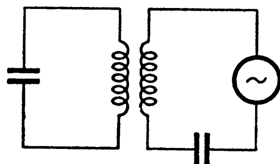
* Wzór (10.28) na pulsację rezonansową w przypadku wymuszonych drgań mechanicznych jest bardziej skomplikowany, gdyż wychylenie punktu drgającego jest analogiem ładunku na okładce kondensatora, natomiast analogiem natężenia prądu jest prędkość tego punktu (przyp Red.).

Amplituda prądu I_{0r} w przypadku rezonansu, jak wynika z (23.15) i (23.16), równa się

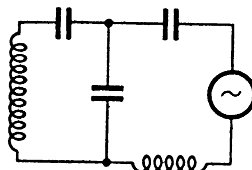
$$I_{0r} = E_0/R,$$

czyli jest wyłącznie funkcją E_0 i R, a nie zależy od L i C.

Zmienność I_0 w zależności od pulsacji Ω wymuszającej SEM dla trzech obwodów o różnych R, a jednakowych L i C, przy stale tej samej wartości E_0 przedstawia rys. 23.6. Widać na nim wyraźnie, że rezonansowy wzrost amplitudy prądu jest tym ostrzej zaznaczony, im mniejsza jest rezystancja obwodu.



Rys. 23.7



Rys. 23.8

Rezonans elektromagnetyczny może także wystąpić podczas oddziaływania na siebie dwóch obwodów elektrycznych LRC o odpowiednio dobranych pulsacjach. Jeden z tych obwodów odgrywa wtedy rolę źródła wymuszającego drgania elektromagnetyczne w drugim obwodzie. Spełnienie warunku rezonansu (23.17) sprowadza się w tym przypadku dc wywołania takiej zmiany pulsacji któregokolwiek z obwodów, by pulsacje ich drgań własnych stały się równe, czyli

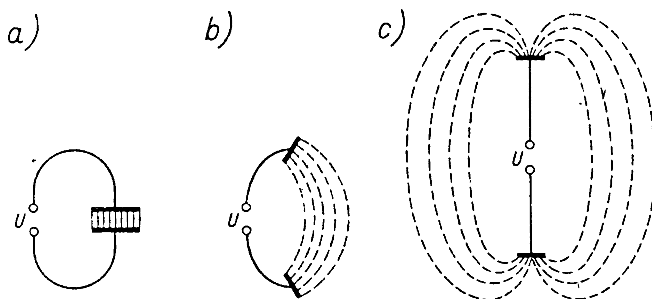
$$L_1 C_1 = L_2 C_2.$$

Wzajemne oddziaływanie obu obwodów może wynikać z ich sprzężenia. Sprzężenia obwodów mogą być różnego typu. Tak np. sprzężenie indukcyjne polega na wzajemnym indukcyjnym oddziaływaniu cewek dwóch obwodów (rys. 23.7), a sprzężenie pojemnościowe na połączeniu dwóch obwodów za pomocą wspólnego kondensatora (rys. 23.8).

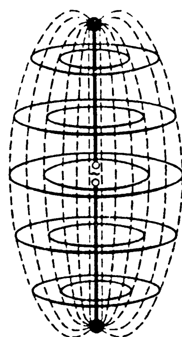
23.4. Elektryczne obwody drgające otwarte

W poprzednich paragrafach tego rozdziału zajmowaliśmy się drganiami elektrycznymi własnymi lub wymuszonymi odbywającymi się w takich obwodach LC lub LRC, które można było nazwać elektrycznymi obwodami drgającymi zamkniętymi ze względu na ograniczenie pola elektrycznego do przestrzeni między okładkami kondensatora, a pola magnetycznego — do obszaru wewnątrz cewki. Obecnie przejdziemy do omówienia zjawisk zachodzących w otoczeniu tzw. elektrycznego obwodu drgającego otwartego, odznaczającego się tym, że powstające pola elektryczne i magnetyczne nie mają ograniczonej przestrzennej lokalizacji. W ten sposób przejdziemy do zjawiska rozchodzenia się w przestrzeni promieniowania elektromagnetycznego, czyli tzw. fali elektromagnetycznej.

Ewolucyjne przejście od obwodu drgającego zamkniętego do obwodu drgającego otwartego przedstawione jest przykładowo na rys. 23.9. Zmiany w stosunku do obwodu zamkniętego np. z rys. 23.1 polegają na usunięciu cewki (indukcyjność obwodu jest teraz związana z pojedynczym przewodem) i stopniowym rozsuwaniu okładek kondensatora. Wspomniane rysunki odnoszą się do warunków elektrostatycznych, wynikających z włączenia do obwodu źródła stałego napięcia. Widać wyraźnie, jak linie sił pola elektrostatycznego w miarę rozsuwania okładek tracą ograniczoną lokalizację między okładkami kondensatora rozprzestrzeniając się coraz szerzej. W omawianych warunkach wobec braku prądu w obwodzie nie ma w jego otoczeniu pola magnetycznego.



Rys. 23.9



Rys. 23.10

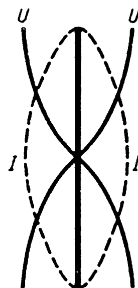
Jeśli jednak w omawianym obwodzie wzbudzimy drgania elektryczne przez sprzężenie go z innym obwodem drgającym lub włączenie generatora drgań, sytuacja ulegnie zmianie. Wystąpienie drgań elektrycznych w takim obwodzie sprowadzi się do pulsowania ładunków od jednej okładki do drugiej, co wywoła powstawanie w otaczającej przestrzeni na przemian zmiennego pola elektrycznego i magnetycznego. Oba pola nie będą już wykazywały ograniczonej lokalizacji. Układ linii sił obu pól w najbliższym otoczeniu rozważanego obwodu przedstawiony jest na rys. 23.10. Linie sił pola elektrycznego (kreskowane) zaczynają się na ładunkach na jednej okładce i kończą na ładunkach na okładce drugiej. Leżą one w płaszczyznach zawierających oś podłużną przewodu łączącego okładki. Linie sił pola magnetycznego (ciągłe) mają postać kół mających środki na wspomnianej osi przewodu i leżących w płaszczyznach prostopadłych do tej osi. Jak się przedstawia rozkład przestrzenny tych linii w większych odległościach od obwodu drgającego, zobaczymy dalej przy bardziej szczegółowej analizie przemian zachodzących podczas drgań elektrycznych w obwodzie tzw. oscylującego dipola elektrycznego.

Najprostszym *oscylującym dipolem elektrycznym* jest pręt z dobrego przewodnika zakończony np. dwiema kulkami, sprzężony w pewien sposób z odpowiednim obwodem drgającym lub połączony z generatorem drgań. Nazwa, jaką nadajemy takiemu obwodowi wymaga pewnych wyjaśnień. Jest w niej wyraźne nawiązanie do znanego nam

już z elektrostatyki pojęcia dipola (por. § 19.2). Dipolem nazywamy układ dwóch ładunków elektrycznych różnoimiennych, równych co do wartości i rozsuniętych na odległość a. Wielkością charakteryzującą dipol jest moment dipolowy $\mathbf{p}$ równy liczbowo iloczynowi Qa, skierowany wzdłuż osi dipola od ładunku ujemnego do dodatniego. W elektrostatyce moment danego dipola na ogół traktowaliśmy jako stały uważając, że mamy do czynienia z dwoma stałymi ładunkami rozsuniętymi na stałą odległość. Zobaczmy, czy stałość momentu cechuje oscylujący dipol elektryczny.

Drgania elektryczne w obwodzie oscylującego dipola sprowadzają się do pulsowania ładunków między kulami. W każdej chwili w całej długości pręta odbywa się ruch elektronów w jednym kierunku (dlaczego?), wskutek czego jedna z kul ładuje się ujemnie (nadmiar elektronów), druga — dodatnio (niedomiar elektronów). Powstaje zatem dipol elektryczny. Zmiany ładunku na kulach mają charakter sinusoidalny, jeżeli taki charakter ma np. wymuszająca siła elektromotoryczna. W ciągu jednego okresu drgań ładunek każdej z kul zmienia się od zera przez maksimum ładunków jednego znaku, spadek do zera, osiągnięcie maksimum ładunków przeciwnego znaku i powrót do zera. Takim sinusoidalnym zmianom ładunków na obu kulach towarzyszy sinusoidalna zmiana momentu dipolowego, jak również sinusoidalna zmiana napięcia między kulami. Zmiany natężenia prądu w obwodzie oscylującego dipola mają również charakter sinusoidalny wykazując jednak przesunięcie fazowe o $\pi/2$ względem ładunku i napięcia (przypominamy zależność $I = dQ/dt$).

Na zakończenie tego krótkiego opisu zmian zachodzących w samym oscylującym dipolu elektrycznym podkreślimy pewną nasuwającą się analogię mechaniczną. Oscylujący dipol można kojarzyć ze stojącą falą napięcia i natężenia prądu. Na końcach dipola mamy pojawiające się co pół okresu maksymalne napięcia, tam więc są zlokalizowane strzałki fali stojącej napięcia. Węzeł tej fali znajduje się w środku dipola. Cała długość dipola odpowiada połowie długości fali. Istnienie strzałek i węzła napięcia można wykazać doświadczalnie za pomocą neonówki* włączonej równolegle na końcach i w środku pręta lub za pomocą innego wskaźnika napięcia.



Rys. 23.11

Falę stojącą napięciową powstającą w oscylującym dipolu przedstawia rys. 23.11. Potencjały poszczególnych punktów dipola są proporcjonalne do odległości tych punktów od nakreślonej linią ciągłą półsinusoidy napięcia. Druga ciągła półsinusoida odpowiada stojącej fali napięcia po upływie połowy okresu. Zaznaczone na tym samym rysunku półsinusoidy kreskowane przedstawiają falę stojącą natężenia prądu: tym razem mamy strzałki w środku dipola, a węzły na końcach. Do doświadczalnego sprawdzenia tych prawidłowości nadaje się słaba żaróweczka wskaźnikowa włączona szeregowo do obwodu.

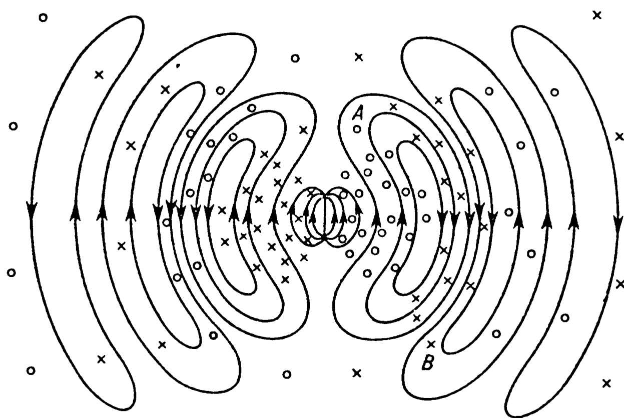
---

* Lampa neonowa (pospolicie zwana neonówką) jest to bańka szklana, wypełniona rozrzedzonym neonem z wtopionymi dwiema elektrodami. Na włączenie stałego napięcia między elektrody neonówka reaguje czerwonym jarzeniem neonu w otoczeniu katody, na napięcie zmienne — o częstotliwości co najmniej kilkunastu herców — jarzeniem w otoczeniu obu elektrod.

W ten sposób zapoznaliśmy się ze zjawiskami zachodzącymi w oscylującym dipolu w czasie drgań elektrycznych. Obecnie przejdziemy do omówienia zachowania się pola elektrycznego i magnetycznego w otoczeniu dipola, tzn. zajmiemy się promieniowaniem dipola.

23.5. Promieniowanie oscylującego dipola elektrycznego

Jak już wspominaliśmy w paragrafie poprzednim, pulsowanie ładunków w dipolu wywołuje powstawanie w jego otoczeniu zmiennego pola elektrycznego i magnetycznego. Schematyczny rozkład linii sił obu pól w najbliższym otoczeniu oscylującego dipola przedstawia rys. 23.10. Układ linii w nieco większej odległości od dipola staje się bardziej skomplikowany, jak to widać z rys. 23.12. Rysunek ten przedstawia przekrój obszaru w pobliżu dipola płaszczyzną zawierającą jego oś. Linie ciągłe przedstawiają linie sił pola elektrycznego, kółka zaś ciemne i krzyżyki są śladami przebicia płaszczyzny rysunku przez linie sił pola magnetycznego odpowiednio wybiegające przed lub przechodzące za płaszczyznę rysunku.



Rys. 23.12

Od razu zwraca uwagę fakt, że linie sił pola elektrycznego leżące nie w bezpośrednim sąsiedztwie dipola nie zaczynają się już i nie kończą na ładunkach istniejących na jego końcach, lecz tworzą pewne krzywe zamknięte. Układ linii sił pola magnetycznego również stwarza pewne niespodzianki. Jak np. tłumaczyć istnienie linii magnetycznych dających ślady takie, jak punkty A lub B, tzn. leżących w płaszczyznach prostopadłych do osi dipola, ale przecinających nie samą oś dipola, lecz jej przedłużenia?

Wspomniane komplikacje wynikają z tego, że zmieniające się w czasie pole elektryczne ma zdolność wywoływania w swoim otoczeniu pola magnetycznego, a zmieniające się w czasie pole magnetyczne ma zdolność wywoływania w swoim otoczeniu pola elektrycznego. Są to konsekwencje podstawowych dla całej elektrodynamiki *praw Maxwella*, o których w tym kursie szerzej nie będziemy mówili. Wspomnimy tylko, że są to cztery równania różniczkowe wiążące elektryczne i magnetyczne pole z ładunkami i prądami,

jak również wzajemnie ze sobą. Niektóre z poznanych przez nas praw, jak prawo Gaussa lub prawo indukcji Faradaya, wynikają bezpośrednio z ogólniejszych od nich praw Maxwella. Na inne wnioski z prawa Maxwella powołamy się jeszcze w § 23.7.

Wróćmy jednak do omawianego przypadku oscylującego dipola elektrycznego. Zgodnie z tym co powiedzieliśmy, gdy do określonego punktu w otoczeniu oscylującego dipola dochodzi zaburzenie w postaci zmiennego pola elektrycznego lub magnetycznego, staje się ono źródłem następujących kolejno po sobie przemian pól (a równocześnie kolejnych przemian energii). Każde z kolejno powstających pól obejmuje już nieco większy obszar przestrzeni. Mamy zatem do czynienia nie tylko z przemianami jednej postaci energii w drugą, lecz równocześnie z ich rozchodzeniem się w przestrzeni. Ponieważ w każdym elemencie przestrzeni istnieje pewne pole elektryczne i magnetyczne związane z odpowiednimi energiami E_e i E_m, występującymi łącznie, więc w tych warunkach uzasadnione jest traktowanie sumarycznej energii jako *energii elektromagnetycznej*. Transport tej energii w otoczeniu oscylującego dipola odpowiada zjawisku *rozchodzenia się promieniowania dipola*, które jest szczególnym przypadkiem *fali elektromagnetycznej*.

Po tych wyjaśnieniach potrafimy już wytłumaczyć istnienie zakwestionowanych poprzednio linii magnetycznych z rys. 23.12, leżących pozornie zbyt wysoko lub zbyt nisko w stosunku do końców osi dipola. Ich istnienie wiąże się już tylko pośrednio z przepływem zmiennego prądu w dipolu; bezpośrednio natomiast jest ono związane ze zmiennym polem elektrycznym rozchodzącej się fali. Podobnie w większych odległościach od dipola pole elektryczne nie jest już bezpośrednio związane z ładunkami na końcach dipola, lecz jest wynikiem istnienia zmiennego pola magnetycznego rozchodzącej się fali.

W naszych rozważaniach doszliśmy do bardzo interesujących wniosków, które zasługują na specjalne podkreślenie: *powstanie fali elektromagnetycznej wymaga istnienia zmiennego ruchu ładunków* (zmiennego prądu), *lecz fala, która już powstała, samej sobie zawdzięcza zdolność rozchodzenia się w przestrzeni — w przypadku braku absorpcji — na nieskończone odległości i w nieograniczonym czasie.* Tak np. fale świetlne (jedna z postaci fali elektromagnetycznej, jak zobaczymy w następnym paragrafie) mogą do nas dochodzić od gwiazd odległych o miliony lat światła po milionach lat od chwili ich wysłania.

Promieniowanie dipola wykorzystamy do ustalenia ważnego powiązania kierunków wektorów **E**, **H** i wektora jednostkowego **k**, który będziemy nazywali wektorem kierunkowym, gdyż reprezentuje on kierunek rozchodzenia się fali elektromagnetycznej. Wspomniane trzy wektory w każdym punkcie przestrzeni są wzajemnie do siebie prostopadłe.

Na zakończenie warto zaznaczyć, że rozkład przestrzenny wypromieniowywanej przez dipol energii nie jest bynajmniej równomierny w jego otoczeniu. Promieniowanie jest maksymalne w kierunku prostopadłym do osi dipola, a równe zeru w kierunku tej osi.

Promieniowaniu modelowego układu oscylującego dipola poświęciliśmy dość dużo uwagi ze względu na to, że wiele rzeczywistych układów elektrycznych promieniujących wykazuje podobne właściwości. Tak np. w liniowej antenie radiowej mamy do

czynienia z pulsowaniem ładunków podobnym do tego, które opisywaliśmy w przypadku oscylującego modelowego dipola. Pola powstające w odległościach wielokrotnie przewyższających długość anteny mają charakter dokładnie taki sam, jak pola w dużych odległościach od oscylującego dipola. Do układów drgających typu dipolowego należą też układy atomowe i cząsteczkowe. Zmianom oscylacji atomów w cząsteczkach, jak też zmianom ruchu obrotowego cząsteczek jako całości, gdy te zmiany są związane ze zmianami momentu dipolowego cząsteczek (por. § 19.16) może towarzyszyć emisja promieniowania elektromagnetycznego o długościach fali od kilku do kilkunastu mikrometrów.

23.6. Fale elektromagnetyczne

Na przykładzie promieniowania oscylującego dipola elektrycznego zapoznaliśmy się już z falami elektromagnetycznymi. Wiemy, że fala elektromagnetyczna sprowadza się do rozchodzenia się w przestrzeni zaburzeń w postaci zmiennych pól elektrycznego i magnetycznego, prostopadłych wzajemnie do siebie i do kierunku ich rozchodzenia się.

Badania doświadczalne fal elektromagnetycznych przyniosły zarówno potwierdzenie hipotezy Maxwella co do ich istnienia, jak też szereg dodatkowych interesujących wiadomości co do ich właściwości. Badania te zostały zapoczątkowane w końcu ub. stulecia przez Hertza, który w warunkach laboratoryjnych, używając pewnej odmiany elektrycznego obwodu drgającego otwartego (zwanej *oscylatorem Hertza*) wytworzył fale elektromagnetyczne, zaliczane dziś do kategorii fal radiowych. W dalszych badaniach udało mu się wykazać, że promieniowanie idące od jego oscylatora 1) ulega odbiciu od powierzchni metalicznych zgodnie z prawami odbicia fal (por. § 11.4), 2) ulega załamaniu przy przejściu z jednego dielektryka do drugiego (np. załamuje się do podstawy w wielkich pryzmatach z parafiny umieszczonych w powietrzu), 3) ulega interferencji, w wyniku której — w odpowiednich warunkach doświadczalnych — może wytwarzać się fala stojąca np. podczas nakładania się fali idącej od źródła i fali odbitej od płyty metalicznej. Przez przesuwanie odpowiednio dobranych i zorientowanych obwodów rezonujących (*rezonatorów Hertza*) między źródłem i płytą odbijającą można ustalać położenie węzłów i strzałek.

Stwierdzenie istnienia interferencji, jak również ugięcia i polaryzacji, a więc zjawisk typowych dla ruchu falowego, potwierdziło naturę falową promieniowania elektromagnetycznego.

23.7. Widmo fal elektromagnetycznych

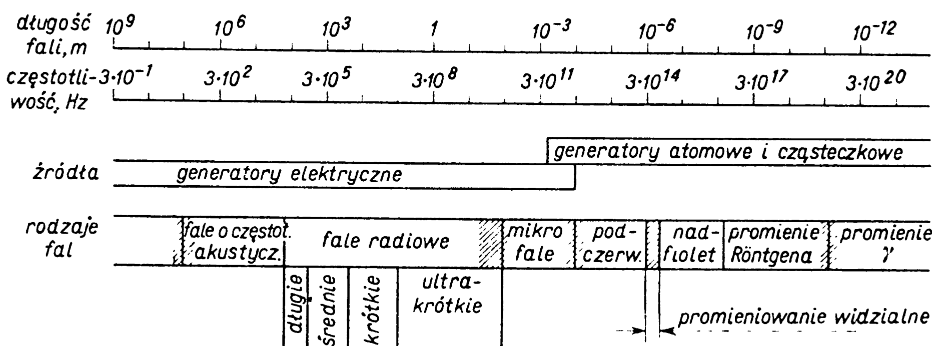
Jednym z wniosków wypływających z rozważań teoretycznych Maxwella było ustalenie, że zarówno *pole elektryczne*, jak i *pole magnetyczne* (*a zatem fala elektromagnetyczna*) *rozchodzą się w próżni z prędkością c równą*:

$$c = \frac{1}{\sqrt{\varepsilon_0 \mu_0}}, \tag{23.19}$$

gdzie ε_0 i μ_0 oznaczają odpowiednio przenikalność elektryczną i przenikalność magnetyczną próżni.

Podstawienie do tego wzoru wartości ε_0 i μ_0 ($\varepsilon_0 = 8{,}840 \cdot 10^{-12}$ C^2/N · m^2; $\mu_0 = 4\pi \cdot 10^{-7}$ Wb/A · m) prowadzi do wartości $c = 2{,}9979 \cdot 10^8$ m/s. Wynik jest bardzo interesujący. Okazuje się, że *prędkość rozchodzenia się fali elektromagnetycznej w próżni jest stała, niezależna od częstotliwości i równa prędkości rozchodzenia się światła w próżni*. To stwierdzenie było zalążkiem hipotezy, z której następnie powstała *teoria elektromagnetyczna światła*. Zgodnie z tą teorią *światło jest jednym z rodzajów promieniowania elektromagnetycznego*.

Dalsze badania, rozwój fizyki współczesnej, jak również rozwój techniki (radio, telewizja, radiolokacja) przyniosły nowe odkrycia w dziedzinie promieniowania elektromagnetycznego. Obecnie wiadomo, że promieniowanie elektromagnetyczne wypełnia w sposób ciągły bardzo szeroki zakres częstotliwości (a więc i długości fal). Ze względu na różne rodzaje stosowanych generatorów i detektorów w różnych przedziałach częstotliwości wygodnie jest wprowadzić pojęcie widma elektromagnetycznego uwzględniające podział promieniowania elektromagnetycznego na różne rodzaje. Trzeba jednak pamiętać, że ostrego podziału między poszczególnymi rodzajami nie ma, że wydzielone obszary widma elektromagnetycznego mniej lub więcej na siebie zachodzą.



Rys. 23.13

Przegląd widma elektromagnetycznego rozpoczniemy od fal najdłuższych, a więc o najmniejszych częstotliwościach (rys. 23.13). Są nimi *najdłuższe fale radiowe* o długościach fal sięgających kilku kilometrów, a częstotliwościach rzędu 150 kHz. Za nimi rozciągają się *fale radiowe średnie* (setki metrów), *krótkie* (dziesiątki metrów) i *ultrakrótkie* (metry, decymetry). Za falami radiowymi rozciąga się obszar *mikrofal*, sięgający do długości fal rzędu 10^{-3} m. Warto wspomnieć o szerokim zastosowaniu fal decymetrowych i centymetrowych w telewizji i radiolokacji. Najkrótsze mikrofale nakładają się na najdłuższe fale z zakresu tzw. *podczerwieni*, tzn. z zakresu promieniowania cieplnego rozciągającego się aż do długofalowej granicy promieniowania tzw. widzialnego.

Zakres *promieniowania widzialnego*, tak niesłychanie ważnego dla człowieka, jest stosunkowo bardzo wąski. Obejmuje on długości fal od $7{,}6 \cdot 10^{-7}$ m do $4 \cdot 10^{-7}$ m. Od strony krótkich fal obszar widzialny graniczy z tzw. nadfioletem (ultrafioletem) sięga-

jącym do długości fal rzędu 10^{-8} m. Tutaj rozpoczyna się dziedzina *promieniowania rentgenowskiego*, pokrywająca się od strony krótkich fal z *promieniowaniem* γ, rozciągającym się do długości fal rzędu 10^{-13}–10^{-14} m (częstotliwość około 10^{22} Hz).

Raz jeszcze podkreślamy, że wszystkie wymienione wyżej rodzaje promieniowania elektromagnetycznego mają tę samą naturę, rozchodzą się w próżni z tą samą prędkością c, różnią się natomiast częstotliwościami i długościami fal. Te różnice decydują o różnych technikach stosowanych do ich wytwarzania, wykrywania i badania ich właściwości. Do wielu z tych spraw powrócimy jeszcze w rozdziałach poświęconych optyce falowej (rozdz. 25 i 26) i fizyce atomowej (rozdz. 30).

Pytania i zadania

1. Jakie procesy nazywamy drganiami elektromagnetycznymi? Opisać najprostszy sposób wzbudzania drgań elektromagnetycznych.

2. Opisać przebieg oscylacyjnego rozładowania kondensatora w obwodzie LC w ciągu jednego okresu z punktu widzenia zmian pól i energii. Wyjaśnić podstawową rolę, jaką w tym procesie odgrywa indukcyjność obwodu.

3. Scharakteryzować stan obwodu LC z rys. 23.1 co $\frac{1}{4}$ okresu z punktu widzenia a) natężenia prądu w obwodzie, b) ładunku na okładce A, c) napięcia między okładkami kondensatora przyjmując, że w chwili $t = 0, I = 0, U = U_0, Q = Q_0$. Które z wymienionych wielkości są w fazie zgodne, a które wykazują przesunięcia fazowe?

4. Jakie rozważania doprowadzają do równania różniczkowego natężenia prądu w przypadku drgań elektrycznych niełumionych w obwodzie LC?

5. Zmiany prądu przy oscylacyjnym rozładowaniu kondensatora mają charakter sinusoidalny. Jaki charakter mają zmiany prądu wynikające z istnienia indukcyjności w obwodzie prądu stałego?

6. W czasie drgań mechanicznych ulegają kolejnym przemianom energia potencjalna i kinetyczna, w czasie drgań elektrycznych — energia elektryczna i magnetyczna. Czy można by i w tym drugim przypadku mówić o przemianach energii potencjalnej i kinetycznej?

7. Jakie zmiany występują w przebiegu drgań elektromagnetycznych, gdy do obwodu pierwotnego LC wprowadzamy rezystancję R?

8. Jakie wielkości charakteryzują szybkość zaniku drgań tłumionych? Podać ich definicje. Powiązać te wielkości z parametrami charakteryzującymi obwód elektryczny. Przeanalizować wpływ zmian R, L i C na szybkość zanikania drgań.

9. W jakich obwodach elektrycznych można zaniedbywać różnicę pulsacji drgań tłumionych i nietłumionych?

10. Jak można wywołać drgania elektromagnetyczne wymuszone? Jakie cechy mają drgania elektromagnetyczne rezonansowe? Jaki jest warunek wystąpienia rezonansu elektromagnetycznego?

11. W jakich obwodach elektrycznych amplituda rezonansowa ma wyraźnie zarysowane maksimum?

12. Wymienić przykłady sprzężeń obwodów elektrycznych. Opisać sposób wywołania rezonansu dwóch obwodów sprzężonych.

13. Wyjaśnić zjawiska zachodzące w: a) oscylującym dipolu elektrycznym, b) jego bliskim otoczeniu, c) dużej odległości od dipola. Podkreślić zasadniczą różnicę w przemianach energetycznych zachodzących w obwodzie LC w czasie rozładowania kondensatora i w otoczeniu oscylującego dipola.

14. Jakie są cechy charakterystyczne fali elektromagnetycznej?

15. Jakie rodzaje fal można wyróżnić w widmie fal elektromagnetycznych?

16. Obwód elektryczny składa się z kondensatora płaskiego (o powierzchni okładek S, odległości między okładkami d i stałej dielektrycznej ośrodka przedzielającego okładki ε) oraz cewki o indukcyjności L. Drgania elektromagnetyczne w obwodzie rozpoczynają się w chwili, gdy napięcie między okład-

kami wynosi U_0. Obliczyć amplitudę prądu, pulsację prądu, okres drgań oraz czas, po upływie którego prąd osiąga maksimum.

$$\text{Odp. } I_0 = \frac{U_0\sqrt{\varepsilon S/d}}{\sqrt{L}}, \quad \omega = \frac{\sqrt{d}}{\sqrt{\varepsilon LS}},$$
$$T = 2\sqrt{\frac{\varepsilon LS}{d}}, \quad t = \frac{\pi}{2}\sqrt{\frac{\varepsilon LS}{d}}.$$

17. Obliczyć czas, po upływie którego amplituda prądu w obwodzie elektrycznym zawierającym pojemność C, indukcyjność L i rezystancję R zmniejszy się a) do połowy, b) e razy (będzie to tzw. stała czasowa obwodu).

$$\text{Odp. } \frac{2L\ln 2}{R}, \quad 2L/R.$$

18. O ile procent pulsacja drgań tłumionych w obwodzie o indukcyjności $L = 10$ mH, pojemności $C = 10^{-6}$ F i rezystancji $R = 0{,}1\ \Omega$ jest mniejsza od pulsacji drgań nietłumionych układu LC o tych samych wartościach L i C?

Odp. 0,01%.

19. Obwód drgający zawiera pojemność C, indukcyjność L i rezystancję R. Znaleźć równanie różniczkowe opisujące zmienność ładunku Q w zależności od czasu podczas rozładowywania kondensatora. Jakie jest rozwiązanie tego równania w przypadku a) ogólnym, b) gdy $R = 0$. Czy parametry charakteryzujące szybkość zaniku prądu i ładunku są jednakowe? Czy z narzuconych warunków początkowych $t = 0$, $U = U_0$ wynikają jednakowe warunki dla I_0 i Q_0? Czy zatem między ładunkiem i prądem istnieje przesunięcie fazowe? Czy to samo dotyczy napięcia i ładunku?

20. W obwodzie radioodbiornika pracującego w pasmie długofalowym ($\lambda \cong 1500$ m) zastosowano kondensator o zmiennej pojemności, przy czym wartość pojemności maksymalnej jest równa $5 \cdot 10^{-10}$ F. Obliczyć indukcyjność zastosowanej cewki oraz ustalić warunek, jaki musi spełniać rezystancja obwodu?

Odp. 1,2 mH, $R < 3000\ \Omega$.

CZĘŚĆ 4

Optyka

Wstęp historyczny — rozwój poglądów na naturę światła

Nazwa „optyka" odnosi się do nauki o promieniowaniu widzialnym i o niektórych rodzajach promieniowania niewidzialnego.

Podstawy naukowe optyki sięgają czasów Newtona, tzn. końca wieku XVII. Newton był twórcą tzw. *teorii korpuskularnej światła*. Według Newtona światło polega na ruchu bardzo drobnych cząstek, zwanych *korpuskułami*. W określonym ośrodku korpuskuły mają ustaloną prędkość i pęd. Na granicy dwóch ośrodków korpuskuły częściowo ulegają odbiciu, częściowo załamaniu. Odbicie korpuskuł odbywa się zgodnie z prawami odbicia dla kul sprężystych. Załamanie wiąże się z tym, że przy przejściu korpuskuł do drugiego ośrodka następuje zmiana ich prędkości wskutek innego oddziaływania cząstek tego drugiego ośrodka. Zmianie prędkości towarzyszy zmiana kierunku ruchu.

Współczesny Newtonowi *Huygens* był twórcą tzw. *teorii falowej światła*. Początkowo teoria Huygensa nie miała licznych zwolenników: być może wpływał na to autorytet Newtona, występującego przeciwko teorii Huygensa. Obie teorie tłumaczyły zjawiska odbicia i załamania; w odniesieniu do zjawisk załamania prowadziły jednak do różnych wniosków. Z teorii Huygensa wynikało, że w ośrodkach takich jak woda światło powinno rozchodzić się z prędkością mniejszą niż w powietrzu. Z teorii Newtona wypływał wniosek przeciwny. Na korzyść jednej lub drugiej teorii mogło przeważyć szalę doświadczenie. Doświadczalne zmierzenie prędkości rozchodzenia się światła w wodzie wykonane zostało dopiero w połowie XIX wieku przez Foucaulta (por. § 24.8). Otrzymane wyniki potwierdziły przewidywania Huygensa.

Teoria falowa nie mogła jednak utrzymać się w takiej postaci, w jakiej podał ją Huygens. Według Huygensa fale świetlne są falami podłużnymi, rozchodzącymi się w ośrodku hipotetycznym, tzw. eterze, któremu trzeba było przypisać dość niezwykłe właściwości mechaniczne w celu wytłumaczenia olbrzymiej prędkości rozchodzenia się światła. Poprawki i uzupełnienia do teorii Huygensa wnieśli m. in. tacy badacze jak Fresnel, Young, Fraunhofer. Im zawdzięczamy np. zbadanie zjawisk ugięcia i interferencji (rozdz. 25). Fresnel na podstawie badań nad polaryzacją (rozdz. 26) doszedł do wniosku, że fale świetlne muszą być falami poprzecznymi.

Dalsze lata wieku XIX przyniosły interesujące odkrycia, wykazujące związek mię-

dzy zjawiskami świetlnymi i zjawiskami magnetycznymi. Między innymi Faraday odkrył skręcenie płaszczyzny polaryzacji w polu magnetycznym.

Z teoretycznych rozważań Maxwell wyciągnął wniosek, że fale elektromagnetyczne rozchodzą się w próżni z prędkością $c = \frac{1}{\sqrt{\varepsilon_0 \mu_0}}$ (wzór (23.19)), co po podstawieniu danych liczbowych daje wartość równą prędkości światła w próżni. Doświadczenia Hertza wykonane kilkanaście lat później potwierdziły teoretyczne przewidywania Maxwella. Na tej podstawie powstała *teoria elektromagnetyczna światła.*

Okres następujący bezpośrednio po doświadczeniach Hertza był okresem triumfu teorii elektromagnetycznej światła. Wkrótce jednak wystąpiły nowe trudności w związku z badaniem rozkładu energii w widmie promieniowania ciała czarnego (§ 29.1). W celu wytłumaczenia rozkładu natężeń promieniowania ciała doskonale czarnego Planck wysunął w roku 1900 hipotezę, że atomy wysyłają energię promienistą nie w sposób ciągły, lecz w postaci określonych porcji, zwanych *kwantami.* Pojedynczy kwant ma energię E proporcjonalną do częstotliwości promieniowania ν:

$$E = h\nu,$$

gdzie h oznacza *stałą Plancka* równą $6{,}625 \cdot 10^{-34}$ J·s.

Dalsze lata przyniosły odkrycie zjawiska fotoelektrycznego, którego wyjaśnienie podał Einstein (r. 1905), wykorzystując, a równocześnie ugruntowując hipotezę kwantów Plancka. Einsteinowi również zawdzięczamy ustalenie związku między energią i masą:

$$E = mc^2,$$

gdzie c oznacza prędkość światła w próżni.

Na tej podstawie każdemu kwantowi energii promieniowania (fotonowi) można przypisać pewną masę i pęd równy $p = h\nu/c$, czyli cechy korpuskularne. W ten sposób w nowej *teorii kwantowej światła* znajdujemy nawiązanie do dawnej teorii korpuskularnej Newtona.

W chwili obecnej istnieje dwoistość opisu zjawisk optycznych, zwana czasem *dualizmem korpuskularno-falowym światła.* Niektóre zjawiska dają się wytłumaczyć zarówno na podstawie teorii falowej, jak i kwantowej, inne tłumaczą się tylko falowo lub tylko kwantowo. Źródło tej dwoistości tkwi w tym, że do opisu zjawisk atomowych, a więc zjawisk mikroświata, stosujemy pojęcia makroskopowe. Do tych spraw wrócimy jeszcze w dalszych rozdziałach.

ROZDZIAŁ 24

Optyka geometryczna

24.1. Pojęcie promienia świetlnego

Podstawowym pojęciem, jakim posługuje się optyka geometryczna, jest pojęcie promienia świetlnego. Przez *promień świetlny* rozumiemy bardzo wąską wiązkę światła, której oś wyznacza kierunek rozchodzenia się energii promienistej. Pojęcie promienia świetlnego wynika bezpośrednio z naszych obserwacji z życia codziennego, które znalazły wyraz w *twierdzeniu o prostoliniowym rozchodzeniu się światła*. Trzeba jednak pamiętać, że jest to tylko twierdzenie przybliżone. Nie stosuje się ono wtedy, gdy światło natrafia na przeszkody (przesłony) lub przechodzi przez małe otwory. W tym przypadku przez „małe" rozumiemy takie otwory, których rozmiary są porównywalne z długością fali światła. Przy przechodzeniu światła przez małe otwory występuje zjawisko ugięcia, które omówimy w § 25.6.

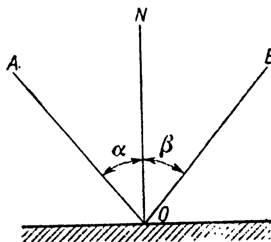
W celu dokładniejszego określenia kierunku rozchodzenia się światła, a więc i kierunku promienia świetlnego, staramy się wyodrębnić możliwie wąskie wiązki świetlne. Doświadczenia jednak wykazują, że nieskończenie wąskich wiązek nie można wyodrębnić wskutek zakłócającego wpływu zjawiska ugięcia (dyfrakcji). Tak więc stwierdzamy, że podstawowe pojęcie optyki geometrycznej, jakim jest pojęcie promienia świetlnego, jest pojęciem fikcyjnym, wynika bowiem z pewnej idealizacji zjawisk. Stąd wniosek, że i wyniki, do jakich dochodzimy w optyce geometrycznej, nie odpowiadają dokładnie rzeczywistości, są tylko pewnym przybliżeniem. W rozległej dziedzinie praktycznych zastosowań optyki przybliżenie to niewiele odbiega od rzeczywistości, można więc szeroko stosować prawa optyki geometrycznej. W przypadku dokładniejszej analizy należy jednak uwzględnić wpływ zjawiska dyfrakcji. Dobry przykład tego znajdziemy w dalszej części tekstu, przy omawianiu teorii mikroskopu początkowo z punktu widzenia optyki geometrycznej, a następnie — z punktu widzenia optyki falowej.

Podstawowymi zjawiskami, jakimi zajmuje się optyka geometryczna, są: odbicie i załamanie światła. Przypomnimy krótko prawa rządzące tymi zjawiskami oraz ich zastosowania.

24.2. Prawa odbicia światła

Doświadczalnie stwierdzono, że odbiciem światła rządzą dwa podstawowe prawa, a mianowicie: 1) *kąt padania* α *i kąt odbicia* β *leżą w jednej płaszczyźnie*, 2) *kąt padania* α *równa się kątowi odbicia* β (rys. 24.1).

Przez *kąt padania* rozumiemy kąt zawarty między promieniem padającym AO i normalną ON, tzn. prostopadłą do powierzchni odbijającej, wystawioną w punkcie padania promienia. Podobnie *kąt odbicia* jest kątem zawartym między normalną i promieniem odbitym OB.



Rys. 24.1

Prawa odbicia są spełnione niezależnie od kształtu powierzchni odbijającej. Te same prawa obowiązują przy odbiciu od powierzchni płaskich, sferycznych, walcowych itp. Po odbiciu od gładkiej powierzchni płaskiej wiązka równoległa pozostaje nadal równoległa. Po odbiciu od powierzchni chropowatej z wiązki promieni równoległych powstaje wiązka promieni biegnących w najrozmaitszych kierunkach. Ten przypadek odbicia nazywamy *rozproszeniem*.

W zależności od rodzaju powierzchni odbijającej (wypolerowana powierzchnia metaliczna, szyba szklana posrebrzona lub pokryta amalgamatem rtęci, powierzchnia wody) mniejszy lub większy procent energii padającego promieniowania ulega odbiciu. Reszta zostaje pochłonięta lub przenika do drugiego ośrodka. Takie powierzchnie, które odbijałyby całkowicie padające promieniowanie, byłyby zwierciadłami idealnymi.

24.3. Zwierciadła płaskie i sferyczne

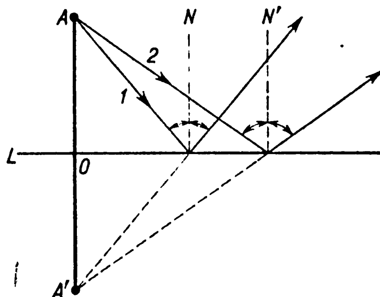
W zależności od kształtu powierzchni odbijającej dzielimy zwierciadła na płaskie, sferyczne (kuliste): wklęsłe i wypukłe, paraboliczne, eliptyczne itp. Przypominamy podstawowe dane dotyczące odbicia promieni od: 1) zwierciadeł płaskich i 2) sferycznych.

1) Załóżmy, że punkt A (rys. 24.2) jest punktowym źródłem światła. L jest powierzchnią zwierciadła płaskiego. Spośród promieni wysyłanych przez źródło wybieramy dwa (zresztą zupełnie dowolne). Stosując do każdego z nich prawa odbicia znajdujemy, że po odbiciu tworzą one wiązkę rozbieżną. Przedłużenia promieni odbitych przecinają się poza zwierciadłem, tworząc w punkcie A' obraz źródła A. *Obraz* ten nazywamy

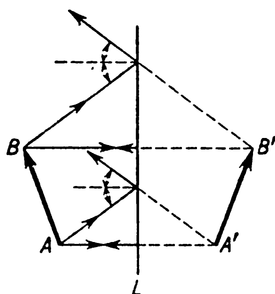
obrazem *pozornym* z tego względu, że powstaje nie wskutek przecięcia rzeczywistych promieni odbitych, lecz wskutek przecięcia przedłużeń tych promieni poza zwierciadło.

Zadanie. Udowodnić, że obraz danego punktu tworzy się w zwierciadle płaskim na przedłużeniu normalnej poprowadzonej z tego punktu do zwierciadła, poza zwierciadłem w takiej odległości, w jakiej przedmiot znajduje się przed zwierciadłem.

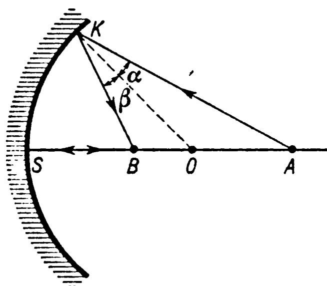
Stosując wniosek z zadania do każdego punktu źródła rozciągłego, np. w postaci strzałki świecącej AB (rys. 24.3), znajdziemy łatwo jej obraz pozorny $A'B'$ tej samej wielkości.



Rys. 24.2



Rys. 24.3



Rys. 24.4

Na podstawie tych pierwszych konstrukcji geometrycznych stwierdzamy, że *odbicie pojedynczego promienia nie może wyznaczyć położenia obrazu.* Do ustalenia położenia punktu A' potrzebne są dwa promienie wychodzące z punktu A, podobnie do wyznaczenia położenia punktu B' potrzebne są dwa promienie wychodzące z punktu B.

2) *Zwierciadla sferyczne* są to takie zwierciadła, których powierzchnia odbijająca stanowi powierzchnię czaszy kulistej. Zwierciadła te dzielą się na: wklęsłe i wypukłe.

Rozpatrzymy początkowo tworzenie się obrazów w *zwierciadle wklęsłym* (rys. 24.4). Wprowadźmy następujące oznaczenia: S — środek czaszy kulistej, O — środek krzywizny zwierciadła, tzn. środek kuli, której część powierzchni stanowi powierzchnię zwierciadła. Prostą przechodzącą przez punkty S i O nazywamy główną osią optyczną

zwierciadła. Odległość $SO = r$ jest promieniem krzywizny zwierciadła, A oznacza punktowe źródło światła.

Zgodnie z tym, co już mówiliśmy, ustalenie położenia obrazu punktu A wymaga zbadania biegu dowolnych dwóch promieni biegnących z punktu A i ulegających odbiciu na powierzchni zwierciadła. Przypuśćmy, że jeden z promieni biegnie wzdłuż osi optycznej. Ponieważ pada on na zwierciadło prostopadle, więc po odbiciu wraca tą samą drogą. Drugi promień AK pada na zwierciadło pod kątem AKO i odbija się pod kątem OKB, równym kątowi OKA. Przecięcie obu promieni odbitych przypada w punkcie B. Punkt ten jest zatem obrazem rzeczywistym źródła A. Łatwo wykazać, że inne promienie wychodzące z punktu A po odbiciu przechodzą również przez punkt B.

Rozważając tworzenie się obrazów w zwierciadłach wklęsłych można w pierwszym przybliżeniu ograniczyć się do promieni tworzących mały kąt z osią główną zwierciadła. Wtedy w przybliżeniu $KA = SA = x =$ odległość przedmiotu od zwierciadła, $KB = SB = y =$ odległość obrazu od zwierciadła.

Na podstawie twierdzenia o dwusiecznej kąta w trójkącie możemy napisać następującą proporcję:

$$\frac{BO}{OA} = \frac{KB}{KA}.$$

Ale $BO = r-y$, $KB = y$,
$OA = x-r$, $KA = x$,

czyli

$$\frac{r-y}{x-r} = \frac{y}{x},$$

$$rx-yx = yx-ry,$$

$$rx+ry = 2yx,$$

$$\frac{1}{x}+\frac{1}{y} = \frac{2}{r}.$$

Oznaczmy $2/r = 1/f$, czyli $f = r/2$. Szukamy znaczenia fizycznego wielkości f. Okazuje się, że gdy $x \to \infty$, to $y \to f$. Innymi słowy, obraz źródła leżącego nieskończenie daleko tworzy się w odległości f od zwierciadła. Tę odległość nazywamy *odległością ogniskową* zwierciadla. Odpowiada ona połowie promienia krzywizny (rys. 24.5).

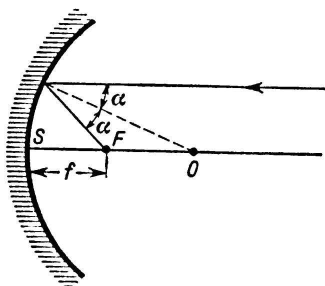
Punkt F na osi głównej, leżący w odległości ogniskowej od zwierciadła, nazywamy *ogniskiem głównym* zwierciadła. Jest to punkt, przez który przechodzą po odbiciu od zwierciadła wszystkie promienie biegnące do zwierciadła równolegle do osi głównej.

Otrzymany wzór

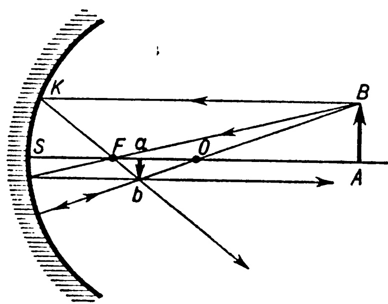
$$\frac{1}{x}+\frac{1}{y} = \frac{1}{f} \tag{24.1}$$

jest podstawowym wzorem wiążącym ze sobą odległości przedmiotu i obrazu od zwierciadła oraz jego odległość ogniskową. Pamiętać musimy, że jest to wzór przybliżony, słuszny dla promieni tworzących mały kąt z osią główną.

W rozważaniach optycznych wykorzystujemy często *odwracalność biegu promieni*: ustawiając źródło na miejscu obrazu otrzymujemy obraz w poprzednim położeniu źródła. Wynika to z praw odbicia, a znajduje wyraz w symetrii wzoru (24.1) względem x i y.



Rys. 24.5



Rys. 24.6

Przypomnimy metodę graficzną stosowaną do przedstawiania tworzenia się obrazów w zwierciadle sferycznym. Niech AB (rys. 24.6) przedstawia przedmiot świecący, SO — oś główną, O — środek krzywizny, F — ognisko główne. W celu znalezienia obrazu dowolnego punktu przedmiotu, np. punktu B, wybieramy takie dwa promienie padające, których bieg po odbiciu jest łatwy do określenia. Może to być promień BK równoległy do osi głównej, który po odbiciu przechodzi przez ognisko. Może to być promień przechodzący przez środek krzywizny O; ponieważ pada on na zwierciadło prostopadle, więc po odbiciu wraca tą samą drogą. Może to być promień biegnący do zwierciadła przez ognisko F, który po odbiciu będzie równoległy do osi głównej itd. Niezależnie od tego, jaką parę promieni wychodzących z punktu B weźmiemy pod uwagę, przetną się one po odbiciu (względnie ich przedłużenia) w tym samym punkcie b, który będzie obrazem rzeczywistym (względnie pozornym) punktu B.

W analogiczny sposób można by znaleźć obraz punktu A leżącego na osi głównej. Położenie obrazu wypada w takim punkcie a, że obraz ab jest prostopadły do osi głównej, jeżeli przedmiot AB jest do osi prostopadły. Dlatego też szukając obrazów przedmiotów prostopadłych do osi wyznacza się tylko położenie obrazu punktu najdalej odsuniętego od osi i następnie prowadzi się z niego linię prostopadłą do osi.

Zadanie. Znaleźć metodą graficzną obrazy wytworzone w zwierciadle wklęsłym przy następujących odległościach strzałki świecącej od zwierciadła: 1) $x > 2f$; 2) $x = 2f$; 3) $f < x < 2f$; 4) $x = f$; 5) $x < f$. Otrzymane wyniki porównać z wynikami podanymi w tabeli 24.1.

Tabela 24.1

Obrazy w zwierciadłach wklęsłych

Odległość przedmiotu od zwierciadła	Odległość obrazu od zwierciadła	Charakterystyka obrazu	Powiększenie W
$x > 2f$	$f < y < 2f$	rzeczywisty, zmniejszony, odwrócony	$0 < W < 1$
$x = 2f$	$y = 2f$	rzeczywisty, równej wielkości, odwrócony	$W = 1$
$f < x < 2f$	$y > 2f$	rzeczywisty, powiększony, odwrócony	$W > 1$
$x = f$	$y = \infty$	—	—
$x < f$	$y < 0$	pozorny, powiększony, prosty	$W < 0, \|W\| > 1$

Wielkość y, odpowiadającą danemu x i f, otrzymujemy przekształcając wzór (24.1):

$$y = \frac{xf}{x-f}. \tag{24.2}$$

Przy pełnej charakterystyce obrazu podajemy zwykle trzy cechy:

1) obraz jest rzeczywisty lub pozorny,
2) prosty lub odwrócony,
3) zwiększony, zmniejszony lub o niezmienionej wielkości.

Ta trzecia cecha wiąże się z pojęciem powiększenia liniowego. Przez *powiększenie liniowe* W rozumiemy stosunek liniowych rozmiarów poprzecznych obrazu do liniowych rozmiarów poprzecznych przedmiotu:

$$W = \frac{ab}{AB}. \tag{24.3}$$

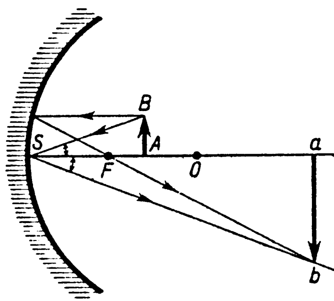
Łatwo udowodnić, że

$$W = \frac{y}{x}. \tag{24.4}$$

Wynika to bezpośrednio z rys. 24.7, na którym do wykreślenia obrazu wykorzystano promień BS. Z podobieństwa trójkątów BSA i abS otrzymujemy proporcję

$$\frac{ab}{AB} = \frac{y}{x}.$$

Powiększenie jest tym większe, im dalej od zwierciadła tworzy się obraz badanego przedmiotu.



Rys. 24.7

Jak widać z tabeli, obraz pozorny jest zawsze prosty. Odpowiada mu powiększenie ujemne o wartości bezwzględnej większej od jedności.

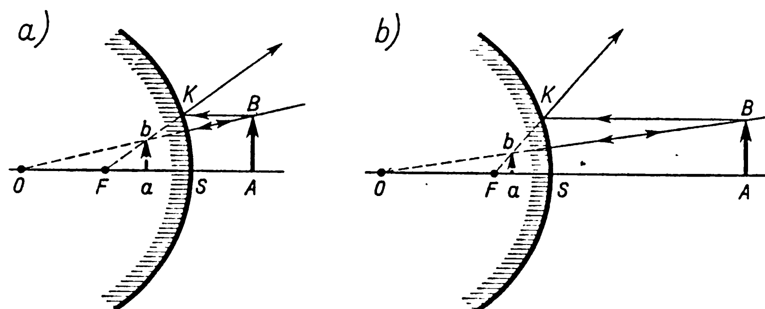
Z tabeli też widać, że w zwierciadłach wklęsłych obraz pozorny powstaje tylko wtedy, gdy x jest mniejsze od f. Inaczej sprawa się przedstawia w przypadku zwierciadeł wypukłych.

W *zwierciadłach wypukłych* promień krzywizny i odległość ogniskową uważamy za wielkości ujemne. Odległość rzeczywistego przedmiotu świecącego od zwierciadła jest zawsze większa od zera.

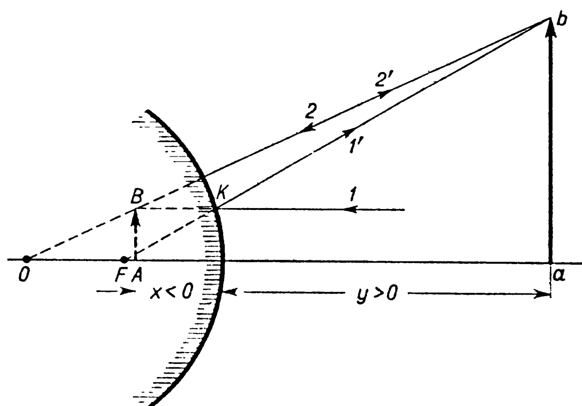
Z wzoru

$$y = \frac{xf}{x-f}$$

wynika, że y jest ujemne. Innymi słowy, obrazy otrzymywane w zwierciadłach wypukłych są zawsze pozorne, niezależnie od odległości przedmiotu świecącego od zwierciadła. Potwierdzają to rys. 24.8a i b.



Rys. 24.8

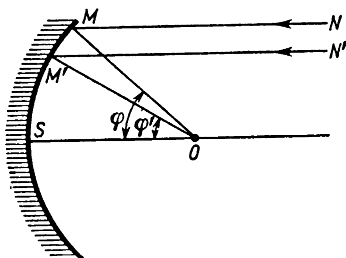


Rys. 24.9

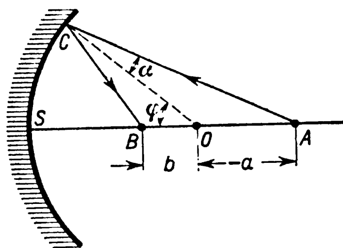
Załóżmy jednak, że na zwierciadło wypukłe padają promienie zbieżne. Gdyby nie było zwierciadła, promienie te tworzyłyby obraz w postaci strzałki AB (rys. 24.9). Aby nie komplikować rysunku, zaznaczamy na nim tylko promienie *1* i *2* jako promienie przyczyniające się do wytworzenia obrazu w postaci punktu B, będącego końcowym punktem strzałki AB. Obraz AB, który wytworzyłby się, gdyby promienie na swej drodze nie trafiały w powierzchnię zwierciadła, można traktować jako przedmiot pozorny. Jego odległość od zwierciadła jest ujemna. Po odbiciu promieni *1* i *2* od zwierciadła wypukłego powstają promienie *1'* i *2'*, wytwarzające obraz rzeczywisty b. Obrazem rzeczywistym całej strzałki AB jest strzałka ab prostopadła do osi głównej zwierciadła.

24.4. Aberracja sferyczna

We wzorze $\frac{1}{x}+\frac{1}{y}=\frac{1}{f}$ nie uwzględniony jest kąt, jaki tworzą promienie padające z osią główną. Pamiętamy jednak, że wzór ten został wyprowadzony dla promieni biegnących blisko osi. Odrzucimy obecnie to założenie i wyprowadzimy ogólniejszy wzór uwzględniający kąt rozwarcia wiązki. Przez *kąt rozwarcia* φ *wiązki* rozumiemy kąt *SOM* (rys. 24.10). Jest to kąt, jaki tworzy oś główna z promieniem krzywizny *OM*, poprowadzonym do skrajnego punktu padania wiązki na zwierciadło. Na przykład dla skrajnego promienia *MN* równoległego do osi kątem rozwarcia będzie właśnie kąt *SOM*. Dla wiązki węższej, ograniczonej promieniem *M'N'*, kąt rozwarcia φ' będzie odpowiednio równy kątowi *M'OS*.



Rys. 24.10



Rys. 24.11

Niech punkt A oznacza punktowe źródło światła, a punkt C — punkt padania na zwierciadło promienia skrajnego (rys. 24.11). Poprowadzenie promienia krzywizny *OC* pozwala ustalić kąt padania α oraz kąt rozwarcia wiązki φ. *CB* oznacza promień odbity. Punkt B jako punkt przecięcia promieni odbitych *CB* i *SA* jest obrazem rzeczywistym punktu A.

Oznaczmy odległość przedmiotu od środka krzywizny zwierciadła przez $-a$ oraz odległość obrazu od środka krzywizny zwierciadła przez b. Zastosujmy twierdzenie sinusów kolejno do trójkątów *OBC* i *OCA*:

$$\frac{b}{\sin\alpha}=\frac{CO}{\sin[180-(\alpha+\varphi)]},$$

$$\frac{-a}{\sin\alpha}=\frac{CO}{\sin(\varphi-\alpha)}.$$

Uwzględniamy, że $CO=r$, odwracamy proporcje i odejmujemy stronami:

$$\frac{\sin\alpha}{b}+\frac{\sin\alpha}{a}=\frac{\sin(\alpha+\varphi)-\sin(\varphi-\alpha)}{r}.$$

Stosujemy wzór na różnicę sinusów

$$\sin\alpha-\sin\beta=2\cos\frac{\alpha+\beta}{2}\sin\frac{\alpha-\beta}{2},$$

skąd

$$\left(\frac{1}{b}+\frac{1}{a}\right)\sin\alpha = \frac{2\cos\varphi\sin\alpha}{r}.$$

Przy założeniu, że $\sin\alpha \neq 0$

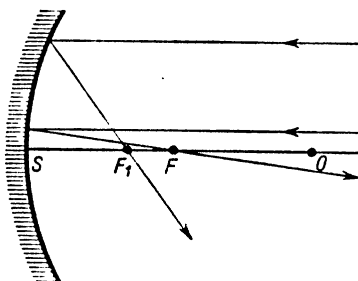
$$\frac{1}{b}+\frac{1}{a} = \frac{2\cos\varphi}{r}. \qquad (24.5)$$

Z równania tego wynika, że dla danego a, tzn. dla danej odległości przedmiotu od środka krzywizny, odległość b obrazu od środka krzywizny jest różna, zależna od kąta rozwartości związanego z badanym promieniem padającym. Innymi słowy, obrazem punktu nie jest jeden punkt. Tę wadę odwzorowań optycznych, związaną z kątem rozwarcia wiązek padających, nazywamy *aberracją sferyczną*.

Załóżmy, że punkt świecący oddala się do nieskończoności; wtedy

$$\frac{1}{b} \to \frac{2\cos\varphi}{r}.$$

W poprzednich naszych rozważaniach, gdy ograniczaliśmy się do bardzo wąskich wiązek padających ($\varphi \to 0$), równoległych do osi głównej, obrazem punktu leżącego w nieskończoności był jeden punkt, zwany ogniskiem głównym. Jego odległość od zwierciadła i od środka krzywizny równała się $r/2$. Badając wiązki o dużym kącie rozwarcia dochodzimy do innego wniosku. Obrazem punktu leżącego w nieskończoności nie jest jeden punkt. Obraz ten jest rozmyty: rozciąga się na osi głównej od odległości $r/2$ do odległości $b' = \dfrac{r}{2\cos\varphi_{max}}$, gdzie φ_{max} oznacza maksymalny kąt rozwarcia wiązki równoległej. Kąt φ_{max} jest zależny od rozmiarów zwierciadła lub od rozmiarów przesłony ograniczającej wiązkę.



Rys. 24.12

Zwracamy uwagę, że wyrażenie $\dfrac{r}{2\cos\varphi_{max}}$ jest większe od $r/2$. Ponieważ odpowiada ono odległości b' między obrazem i środkiem krzywizny, przeto ognisko F_1 skrajnych promieni równoległych leży bliżej zwierciadła niż ognisko F dla promieni bliskich osi. Odległość ogniska F od ogniska F_1 (rys. 24.12) nosi nazwę *aberracji sferycznej podłuż-*

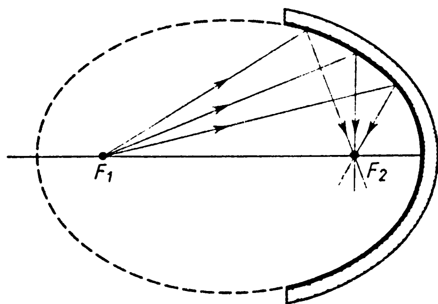
nej. Ustawiając ekran prostopadle do osi głównej między F i F_1 otrzymujemy obraz punktu świecącego nieskończenie odległego nie w postaci punktu, lecz w postaci okrągłej plamki, zwanej *kolem rozproszenia*. Podczas przesuwania ekranu od F do F_1 zmienia się promień koła. Najmniejsza wartość promienia koła rozproszenia nosi nazwę *aberracji sferycznej poprzecznej*.

24.5. Astygmatyzm

Pojęcie ogniska głównego odnosi się do wiązki promieni równoległych do osi głównej. Promienie tworzące wiązki równoległe nachylone pod niewielkimi kątami do osi głównej przecinają się po odbiciu od zwierciadła w punkcie leżącym nie na osi głównej, lecz z boku. W tych warunkach mówimy o ogniskach bocznych albo pochylonych. Inaczej przedstawia się wynik odbicia, gdy mamy do czynienia z wiązkami równoległymi, tworzącymi duże kąty z osią główną. W tym przypadku znowu nie otrzymujemy punktowego odwzorowania punktu świecącego. Zamiast obrazu punktowego otrzymujemy obraz w postaci dwóch wichrowatych odcinków. Tę wadę odwzorowań optycznych nazywamy *astygmatyzmem*. Odległość obu wichrowatych odcinków jest miarą astygmatyzmu wiązki odbitej.

24.6. Zwierciadła niesferyczne

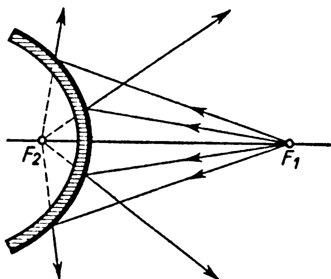
Stosując zwierciadła paraboliczne, eliptyczne lub hiperboliczne można w przypadku pewnych ustawień punktowego źródła otrzymać dokładnie punktowy jego obraz. Innymi słowy, w zwierciadłach tego typu możliwe jest otrzymanie obrazu bez aberracji sferycznej. Wynika to z właściwości geometrycznych tego rodzaju powierzchni. Elipsoidy, paraboloidy i hiperboloidy mają tę właściwość, że normalne wystawione do ich



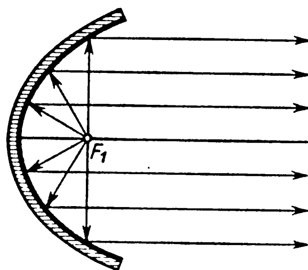
Rys. 24.13

powierzchni w dowolnym punkcie są dwusiecznymi kątów utworzonych przez promienie biegnące od ognisk. Oba ogniska elipsoidy są rzeczywiste i leżą w skończonej odległości. Wszystkie promienie wychodzące z jednego ogniska F_1 przecinają się po odbiciu od zwierciadła eliptycznego w jego ognisku drugim F_2 (rys. 24.13).

W przypadku zwierciadła hiperbolicznego jedno z ognisk jest pozorne. Promienie wychodzące z ogniska F_1 po odbiciu od zwierciadła hiperbolicznego biegną tak, jak gdyby wychodziły z ogniska F_2 (rys. 24.14). Punkt F_2 jest więc obrazem pozornym punktu F_1.



Rys. 24.14

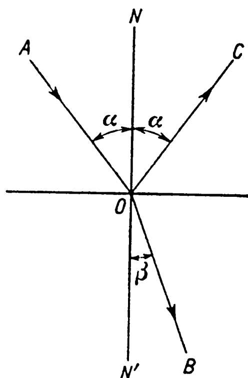


Rys. 24.15

W zwierciadłach parabolicznych jedno z ognisk leży w nieskończoności. Ustawiając punktowe źródło światła w ognisku F_1 (rys. 24.15) otrzymamy po odbiciu wiązkę dokładnie równoległą. I odwrotnie, wiązka równoległa padająca na zwierciadło paraboliczne przekształca się po odbiciu na wiązkę zbieżną, przecinającą się w ognisku F_1. Zwierciadła paraboliczne nie wykazują aberracji sferycznej dla promieni równoległych, nawet w przypadku dowolnie dużych rozmiarów wiązek. Zwierciadła te znalazły zastosowanie w reflektorach.

24.7. Prawa załamania światła

Zazwyczaj przy padaniu światła na granicę dwóch ośrodków przezroczystych, różniących się prędkością rozchodzenia się światła, część energii promienistej przechodzi do drugiego ośrodka w postaci promienia załamanego. Załamanie światła na granicy dwóch ośrodków izotropowych podlega następującym prawom:



Rys. 24.16

1. *Kąt padania α i kąt załamania β leżą w jednej płaszczyźnie. Kąt załamania* jest to kąt utworzony przez normalną ON' i promień załamany OB (rys. 24.16).

2. *Stosunek sinusa kąta padania do sinusa kąta załamania równa się stosunkowi prędkości v_1 rozchodzenia się światła w ośrodku pierwszym do prędkości v_2 rozchodzenia się światła w ośrodku drugim.* Jest to wielkość stała dla danych dwóch ośrodków i danego rodzaju promieniowania

$$n_{2/1} = \frac{\sin\alpha}{\sin\beta} = \frac{v_1}{v_2}. \tag{24.6}$$

Wielkość $n_{2/1}$ nazywamy *współczynnikiem załamania ośrodka drugiego względem pierwszego.* Wprowadzamy też pojęcie *bezwzględnego współczynnika załamania.* Jest to współczynnik za-

łamania danego ośrodka względem próżni. W próżni wszystkie rodzaje promieniowania rozchodzą się z tą samą prędkością. Prędkość tę wyznaczano różnymi metodami. Parę podstawowych metod omówimy w następnym paragrafie.

Oznaczmy bezwzględne współczynniki załamania dwóch ośrodków, np. ośrodka A i ośrodka B, przez n_A i n_B:

$$n_A = \frac{c}{v_A}, \quad n_B = \frac{c}{v_B},$$

skąd względny współczynnik załamania ośrodka B względem A (przejście promienia z A do B)

$$n_{B/A} = \frac{v_A}{v_B} = \frac{n_B}{n_A}.$$

W tabeli 24.2 podane są wartości bezwzględnych współczynników załamania różnych ośrodków, obliczone dla żółtej linii sodu, tzw. linii D, której długość fali wynosi około 5893 Å.

Tabela 24.2

Bezwzględne współczynniki załamania światła

alkohol	1,369	szkło ołowiowe (flint)	1,743
azot	1,000298	szkło potasowe (crown)	1,515
diament	2,41	tlen	1,000271
powietrze suche	1,0002929	woda	1,333
sól kuchenna	1,544	wodór	1,000139
szkło kwarcowe	1,46		

Jak widać z tabeli, współczynnik załamania gazów bardzo mało różni się od jedności, tzn. prędkość rozchodzenia się światła w ośrodkach gazowych mało różni się od prędkości światła w próżni. Przybliżona wartość prędkości światła w próżni (i w powietrzu) wynosi 300 milionów metrów na sekundę.

Podstawowe prawa optyki geometrycznej, takie jak prawo odbicia $\alpha_{pad} = \beta_{odb}$ i prawo załamania $\sin\alpha/\sin\beta = n_{2/1}$ można udowodnić opierając się na *zasadzie Fermata* sformułowanej już w XVII wieku. Zgodnie z tą zasadą *promień świetlny biegnie od dowolnego punktu A do innego punktu B taką drogą, by czas potrzebny na jej przebycie był ekstremalny, tzn. najkrótszy lub najdłuższy z możliwych.* Zazwyczaj jest to czas najkrótszy. Można też mówić o wyborze takiego toru, wzdłuż którego droga optyczna promienia jest ekstremalna. Przez *drogę optyczną* promienia w ośrodku o współczynniku załamania n rozumiemy iloczyn drogi rzeczywistej przebytej w tym ośrodku i współczynnika załamania n. W ten sposób przeliczamy drogę przebytą w danym ośrodku na drogę przebytą w tym samym czasie w próżni.

Zadanie. Udowodnić równość kątów padania i odbicia oraz zależność $\sin\alpha/\sin\beta = n$ opierając się na zasadzie Fermata (por. zadanie 22).

24.8. Prędkość światła

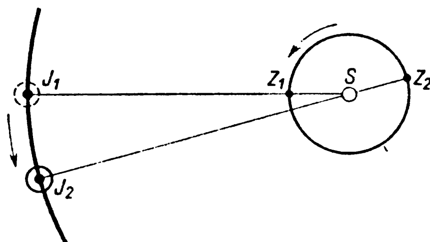
Po raz pierwszy prędkość światła wyznaczona została w roku 1675 przez Olafa Römera.

Römer oparł swoje rozważania na wynikach badania zaćmień jednego z księżyców Jowisza. Gdyby Ziemia i Jowisz zajmowały niezmienne położenia względem siebie i względem Słońca, to okresy czasu upływające między początkami kolejnych zaćmień danego księżyca byłyby niezmienne. Tymczasem badania wykazały, że w ciągu roku okresy te ulegają systematycznym zmianom, np. od wielkości minimalnej do maksymalnej i z powrotem do minimum. Zmiany obserwowanych okresów Römer kojarzył z tym, że wskutek ruchów obu planet dokoła Słońca zmieniają się ich wzajemne odległości. Na uproszczonym rys. 24.17, gdzie orbity obu planet przedstawione są jako koła, mamy zaznaczone położenia Z_1 i J_1, gdy obie planety są najbliżej siebie, i położenia Z_2 i J_2, gdy ich wzajemna odległość jest maksymalna. Różnica obu odległości odpowiada średnicy orbity Ziemi, czyli $2R_Z$. Różnica czasów trwania najdłuższego i najkrótszego zaćmienia stanowi czas potrzebny do przebycia drogi $2R_Z$ przez promień świetlny. Z wzoru

$$v = \frac{s}{t}$$

Römer otrzymał przybliżoną wartość prędkości rozchodzenia się światła w próżni około $2 \cdot 10^8$ m/s.

Pomiary bezpośrednie prędkości światła wykonane zostały po raz pierwszy na Ziemi w połowie XIX wieku. Opracowane zostały dwie metody: metoda Fizeau i metoda Foucaulta. Metoda Fizeau wykorzystuje obrót koła zębatego, a metoda Foucaulta — obrót zwierciadła płaskiego. W metodzie pierwszej potrzebna jest odległość kilku kilometrów, metoda druga daje dobre wyniki przy odległościach około 20 metrów, a więc w warunkach laboratoryjnych.

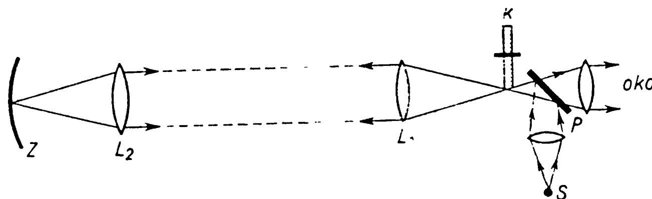


Rys. 24.17

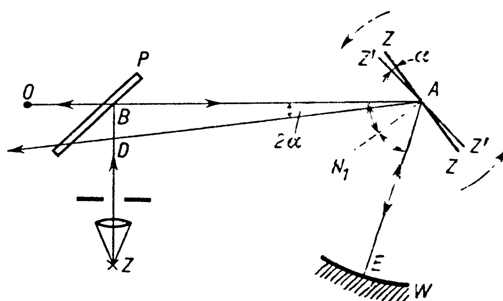
Zasada metody Fizeau przedstawiona jest na rys. 24.18. Promienie biegnące od źródła S po załamaniu w soczewce odbijają się od ukośnie ustawionej płytki szklanej P i tworząc wiązkę zbieżną trafiają na obwód obracającego się koła zębatego K. Gdy punkt zbieżności tej wiązki trafia na otwór między zębami, promienie przechodzą dalej przez układ dwóch achromatycznych soczewek L_1 i L_2 do zwierciadła wklęsłego Z ustawionego tak, że jego środek krzywizny przypada w środku soczewki L_2. Po odbiciu promienie wracają tą samą drogą. Jeśli prędkość obrotu koła zębatego jest tak dobrana, że wracające promienie trafiają na „ząb" koła, a nie na wolną przestrzeń między zębami — oko nie dostrzega źródła. Jeżeli taka sytuacja zachodzi po raz pierwszy przy częstotliwości obrotu ν, to powtórzy się przy częstotliwościach będących całkowitymi wielokrotnościami tej częstotliwości minimalnej. Łatwo sprawdzić, że jeśli n oznacza liczbę zębów na obwodzie koła, ν — minimalną częstotliwość obrotów, przy której źródło znika z pola widzenia, a L — odległość od zwierciadła wklęsłego do koła zębatego, to prędkość światła

$$c = \frac{2L}{t} = \frac{2L}{\frac{1}{2\nu n}} = 4Ln\nu.$$

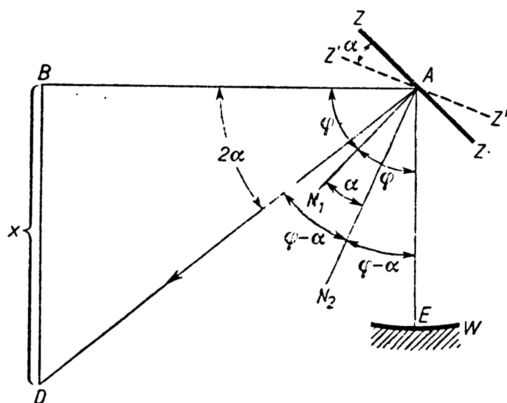
Zasada metody Foucaulta przedstawiona jest na rys. 24.19. Z promieniowania biegnącego od silnego źródła Z wyodrębniamy za pomocą przesłony wąską wiązkę, którą rzucamy na ustawioną ukośnie płytkę szklaną P. Po odbiciu od płytki wiązka pada na zwierciadło płaskie w punkcie A i po odbiciu od niego zostaje skierowana na zwierciadło wklęsłe W. To ostatnie jest tak ustawione, że w punkcie A przypada jego środek krzywizny. A zatem promień padający wzdłuż AE wraca po odbiciu tą samą drogą do



Rys. 24.18



Rys. 24.19



Rys. 24.20

zwierciadła płaskiego i dalej do płytki P, przenika przez nią i trafia do oka obserwatora O. Jeżeli jednak zwierciadło płaskie wykonuje ruch obrotowy, to promień świetlny po dwukrotnym przejściu drogi AE padnie na zwierciadło obrócone o kąt α. Jak łatwo sprawdzić (korzystając z pomocniczego, zrobionego w powiększonej skali rys. 24.20),

nowy promień odbity od obróconego zwierciadełka (linia kreskowana $Z'Z'$) odchylony jest od pierwotnego promienia padającego BA o kąt 2α. Ponieważ w doświadczeniu Foucaulta kąt ten jest niewielki, więc nie popełniamy większego błędu przyjmując

$$2\alpha \approx \operatorname{tg} 2\alpha = \frac{\text{przesunięcie } BD}{\text{odległość } AB} = \frac{x}{L},$$

skąd

$$\alpha = \frac{x}{2L}.$$

Znajomość kąta α umożliwia obliczenie krótkiego czasu, w ciągu którego zwierciadło obróciło się o kąt α. Jeśli bowiem częstotliwość obrotów zwierciadła wynosi ν, to obrót o kąt α odbywa się w czasie

$$t = \frac{1}{2\pi\nu} \frac{x}{2L}.$$

A zatem prędkość światła c, równa stosunkowi przebytej drogi $2AE = 2l$ do czasu, w którym ta droga została przebyta, może być obliczona w metodzie Foucaulta z wzoru

$$c = \frac{2l}{t} = \frac{8\pi\nu Ll}{x}.$$

Wielkość przesunięcia x w pomiarach Foucaulta wynosiła zaledwie 0,7 mm przy odległościach $l = 20$ m. Wielkość tę mierzono za pomocą okularu mikrometrycznego z dokładnością do 0,005 mm, otrzymując dość dokładne wyniki. Innymi słowy posługiwano się mikroskopem, którego okular zaopatrzony był w nić przesuwaną przez pole widzenia za pośrednictwem śruby mikrometrycznej. Metoda Foucaulta, w której do pomiaru prędkości światła wystarczają stosunkowo niewielkie odległości, rzędu kilku do kilkunastu metrów, umożliwiła pomiar c nie tylko w powietrzu, lecz i w innych ośrodkach, np. w wodzie, w dwusiarczku węgla itp. Promień świetlny przebywał drogę AE w rurze wypełnionej badaną substancją.

Po pomiarach przeprowadzonych przez Fizeau i Foucaulta liczni fizycy zajmowali się wyznaczaniem prędkości rozchodzenia się światła w próżni i w powietrzu. Warto tu wspomnieć nazwisko Michelsona, który właściwie całe życie poświęcił na ulepszanie swych metod. W tabeli 24.3 podane są wyniki kilku pomiarów prędkości światła wykonanych różnymi metodami.

Na podstawie wyników wielu pomiarów przyjmuje się obecnie wartość prędkości światła w próżni

$$c = (2{,}997924 \pm 0{,}000010) \cdot 10^8 \text{ m/s}$$

jako najbardziej zbliżoną do rzeczywistej.

24.9. Zjawisko całkowitego wewnętrznego odbicia

Rozważmy przejście promienia z ośrodka B o mniejszej prędkości rozchodzenia się światła do ośrodka A o prędkości większej (rys. 24.21).

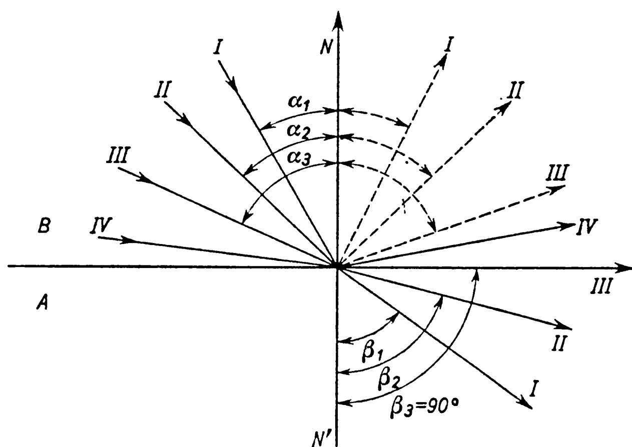
Promień padający I pada pod kątem α_1, załamuje się pod kątem β_1, a równocześnie

Tabela 24.3

Wyniki kilku pomiarów prędkości światła

Nazwisko	Data pomiaru	Wynik m/s	Błąd m/s
Fizeau	1849	$3{,}13300 \cdot 10^8$	—
Foucault	1862	$2{,}98000 \cdot 10^8$	500000
Michelson	1880	$2{,}99910 \cdot 10^8$	50000
Michelson	1926	$2{,}99796 \cdot 10^8$	4000
Anderson	1941	$2{,}99776 \cdot 10^8$	14000
Bergstrand	1950	$2{,}997927 \cdot 10^8$	250
Edge	1956	$2{,}997929 \cdot 10^8$	200

częściowo się odbija, co jest zaznaczone na rysunku linią kreskowaną. Kąt $\beta_1 > \alpha_1$, ponieważ $v_B < v_A$. Promień *II* pada pod kątem α_2, załamuje się pod kątem β_2 większym od α_2. Dalszy promień *III* może padać pod takim kątem α_3, któremu odpowiada kąt załamania β_3 równy 90°. Jeśli jeszcze bardziej powiększymy kąt padania promienia, to już nie otrzymamy promienia załamanego: cała energia promienia padającego przypadnie na promień odbity. W tych warunkach mówimy o *zjawisku całkowitego wewnętrznego odbicia.*



Rys. 24.21

Ustalamy warunki, jakie muszą być spełnione, aby mogło wystąpić zjawisko całkowitego wewnętrznego odbicia. Przede wszystkim promień musi przechodzić z ośrodka *B* o mniejszej prędkości rozchodzenia się światła (czyli z ośrodka *optycznie gęstszego*) do ośrodka *A* o prędkości większej (*optycznie rzadszego*). Poza tym kąt padania promienia musi być większy od tzw. kąta granicznego. Przez *kąt graniczny* rozumiemy taki kąt padania, któremu odpowiada kąt załamania równy 90°. Na rysunku 24.21 kąt α_3 jest właśnie kątem granicznym. Kąt ten spełnia równanie

$$\frac{\sin \alpha_{gran}}{\sin 90°} = n_{A/B} = \frac{1}{n_{B/A}}, \text{ czyli } \sin \alpha_{gran} = \frac{n_A}{n_B},$$

gdzie n_A i n_B są odpowiednio bezwzględnymi współczynnikami załamania ośrodka optycznie rzadszego A i gęstszego B.

Tabela 24.4

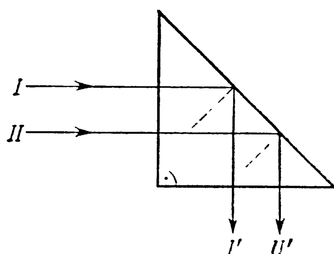
Kąty graniczne (względem powietrza)

woda	około 48°
szkło potasowe (crown)	około 42°
szkło ołowiowe (flint)	około 33°
diament	około 24°

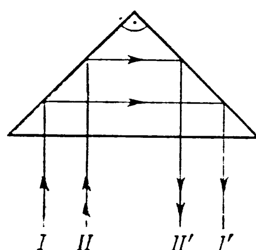
W tabeli 24.4 podane są wartości kątów granicznych dla kilku substancji.

Stosunkowo mały jest kąt graniczny diamentu. Dzięki temu przy odpowiednim oszlifowaniu ścianek (brylant) otrzymujemy na nich wielokrotne całkowite odbicie padających promieni, powodujące piękny połysk brylantu.

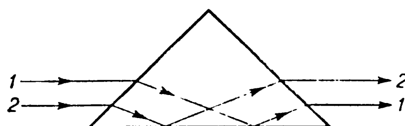
Zjawisko całkowitego wewnętrznego odbicia w równoramiennych prostokątnych pryzmatach szklanych przedstawiono na rys. 24.22, 24.23 i 24.24.



Rys. 24.22



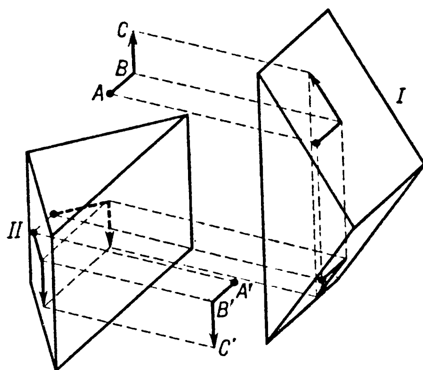
Rys. 24.23



Rys. 24.24

Wiązka równoległa, padając prostopadle na ściankę pryzmatu, nie ulega załamaniu. Na drugą ściankę promienie padają pod kątem 45°, czyli pod kątem większym od granicznego (tab. 24.4). Pryzmat ustawiony względem wiązki padającej tak, jak na rys. 24.22, zmienia kierunek biegu promieni o 90°: przy ustawieniu takim, jak na rys. 24.23 (pryzmat Porro) lub na rys. 24.24 (pryzmat Dovego), otrzymujemy odwrócenie: promień

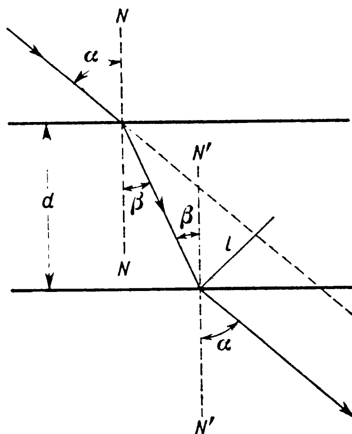
lewy z wiązki padającej przekształca się na promień prawy w wiązce odbitej. Zestawienie takich dwóch pryzmatów służy do usuwania odwrócenia obrazów. Jako przykład może służyć zestawienie pryzmatów w lornetce pryzmatycznej, której schemat przedstawiony jest na rys. 24.25. Jeden z pryzmatów zmienia promień lewy na prawy, a drugi — promień dolny na górny.



Rys. 24.25

24.10. Załamanie światła w płytce płasko-równoległej

Podczas przejścia przez płytkę o ściankach płaskich i równoległych promień ulega dwukrotnemu załamaniu (rys. 24.26). Jeśli ośrodek, z którego promień wchodzi do płytki i ośrodek, do którego on z płytki wychodzi, jest ten sam, promień ulega równoległemu przesunięciu.



Rys. 24.26

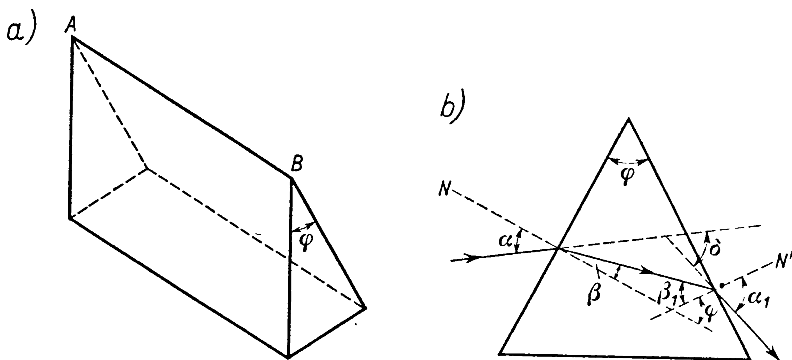
Zadanie. Na podstawie rys. 24.26 udowodnić, że wartość równoległego przesunięcia promienia przy przejściu przez płytkę płasko-równoległą wynosi

$$l = d\sin\alpha\left(1-\sqrt{\frac{\cos^2\alpha}{n^2-\sin^2\alpha}}\right),$$

gdzie d oznacza grubość płytki, α — kąt padania promienia na płytkę, n — współczynnik załamania płytki względem otaczającego ośrodka.

24.11. Załamanie światła w pryzmacie

Bryła ograniczona ścianami płaskimi przecinającymi się pod kątem φ stanowi pryzmat (rys. 24.27a). Kąt dwuścienny φ nosi nazwę *kąta łamiącego* pryzmatu. Prosta AB, wzdłuż której przecinają się płaszczyzny ścian bocznych, nosi nazwę *krawędzi pryzmatu*. Ściana przeciwległa kątowi φ może mieć kształt dowolny, gdyż nie ma ona wpływu na bieg promieni. Często na schematycznych rysunkach pryzmatów trzecia ściana przecina ściany boczne w równych odległościach od krawędzi. Wtedy pryzmat w przekroju prostopadłym do krawędzi przedstawia się jako trójkąt równoramienny (rys. 24.27b). Ściana przeciwległa do krawędzi nosi nazwę *podstawy pryzmatu.*



Rys. 24.27

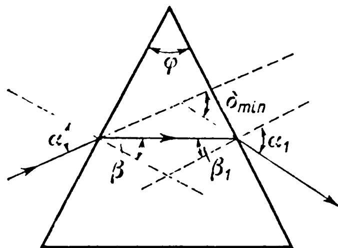
Niech promień monochromatyczny (jednobarwny) pada na ścianę pryzmatu pod kątem α, załamuje się pod kątem β, pada na drugą ścianę pod kątem β_1 i wychodzi z niej pod kątem α_1. Normalne do obu ścian bocznych pryzmatu tworzą ze sobą kąt równy kątowi łamiącemu φ. Przedłużenia promieni padającego i wychodzącego z pryzmatu tworzą kąt δ, zwany kątem odchylenia.

Postaramy się wyrazić współczynnik załamania pryzmatu w zależności od kątów φ i δ. W tym celu musimy znaleźć zależności między kątami: φ, δ, α, i β, aby móc podstawić odpowiednie wartości do wzoru na współczynnik załamania $n = \frac{\sin\alpha}{\sin\beta}$. Z twierdzenia o kącie zewnętrznym w trójkącie wynika, że

$$\varphi = \beta+\beta_1, \qquad \delta = \alpha-\beta+\alpha_1-\beta_1.$$

We wzorach tych wystąpiły dodatkowo kąty α_1 i β_1. Doświadczenie uczy, że zmiana

kąta padania α pociąga za sobą zmianę kąta α_1 i zmianę kąta δ. Początkowo przy zmniejszaniu kąta padania α kąt odchylenia δ stopniowo się zmniejsza, przy pewnej wartości kąta α osiąga wartość minimum, a następnie przy dalszym zmniejszaniu kąta α kąt odchylenia znów rośnie. Tej najmniejszej wartości kąta odchylenia, zwanej kątem najmniejszego odchylenia δ_{min}, odpowiada równość kątów α i α_1. Stąd wynika też równość kątów β i β_1. Promień załamany biegnie w tych warunkach równolegle do podstawy pryzmatu (rys. 24.28).



Rys. 24.28

Gdy spełnione są omówione wyżej zależności, równania ostatnie przyjmują postać

$$\varphi = 2\beta, \qquad \delta_{min} = 2\alpha - 2\beta.$$

A zatem

$$\beta = \frac{\varphi}{2}, \quad \alpha = \frac{\delta_{min}+\varphi}{2}.$$

Po podstawieniu do wzoru na współczynnik załamania otrzymujemy

$$n = \frac{\sin\dfrac{\delta_{min}+\varphi}{2}}{\sin\dfrac{\varphi}{2}}. \tag{24.7}$$

Kąt φ zależny od kształtu pryzmatu jest dla danego pryzmatu stały. Kąt δ_{min} natomiast jest różny dla promieni różnych barw. Gdy na pryzmat pada wiązka promieniowania złożonego (niemonochromatycznego), to każdy ze składowych promieni załamuje się pod innym kątem. W tych warunkach mówimy o rozszczepieniu światła w pryzmacie (§ 28.1 i 2). Jeśli na pryzmat szklany pada wąska wiązka światła białego, to na ekranie ustawionym poza pryzmatem otrzymamy smugę świetlną, w której barwy przechodzą jedna w drugą w sposób ciągły. Smugę tę nazywamy widmem ciągłym. Najbardziej odchylony od pierwotnego biegu jest promień fioletowy, najmniej odchylony — czerwony. Jest to układ promieni w widmie przeciwny do tego, który występuje przy ugięciu w siatce dyfrakcyjnej (por. punkt 25.6.4).

Inną zależność między wartością współczynnika załamania, kątem odchylenia i kątem łamiącym znajdziemy przy założeniu, że kąt łamiący pryzmatu i kąt padania promienia na pryzmat są bardzo małe. Wtedy opierając się na przybliżonej równości $\sin\alpha \cong \alpha$ i $\sin\beta \cong \beta$ piszemy

$$n = \frac{\alpha}{\beta}, \quad n = \frac{\alpha_1}{\beta_1},$$

skąd

$$\alpha = n\beta, \quad \alpha_1 = n\beta_1.$$

Podstawiając powyższe wartości do wzoru na kąt odchylenia δ otrzymujemy

$$\delta = \alpha+\alpha_1-(\beta+\beta_1),$$

$$\delta = n\beta+n\beta_1-\beta-\beta_1,$$

$$\delta = (n-1)(\beta+\beta_1),$$

ale

$$\beta+\beta_1 = \varphi,$$

więc

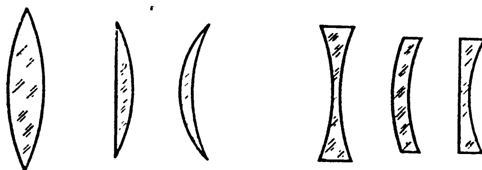
$$\delta = (n-1)\varphi. \tag{24.8}$$

Zależność tę będziemy wykorzystywali przy wyprowadzaniu wzoru dla cienkich soczewek.

24.12. Soczewki cienkie

Dotychczas rozważaliśmy załamanie światła na powierzchniach płaskich, obecnie przechodzimy do badania załamania światła w soczewkach.

Soczewkami nazywamy bryły przezroczyste, ograniczone powierzchniami regularnymi: kulistymi, parabolicznymi, walcowymi itp. Zajmiemy się soczewkami sporządzonymi z materiału optycznie gęstszego niż otaczający ośrodek. Tego rodzaju soczewki działają skupiająco na wiązkę promieni, jeśli w środkowej części są grubsze niż na brzegach. Na rys. 24.29 przedstawione są trzy rodzaje soczewek kulistych skupiających i trzy rodzaje soczewek kulistych rozpraszających. Są to soczewki dwuwypukłe, płasko-wypukłe i wklęsło-wypukłe oraz soczewki dwuwklęsłe, płasko-wklęsłe i wypukło-wklęsłe. Soczewki zbierające przyjęto oznaczać schematycznie na rysunkach symbolem ↕, a soczewki rozpraszające — symbolem ⨶.



Rys. 24.29

Podobnie jak w przypadku zwierciadeł sferycznych wprowadzamy w odniesieniu do soczewek pojęcie środka krzywizny i osi głównej. Przez r_1 i r_2 oznaczamy promienie krzywizn obu powierzchni soczewek. Jeśli grubość soczewki jest mała w porównaniu z promieniami krzywizny, mówimy o *soczewkach cienkich.*

Zajmiemy się teorią soczewek cienkich, ograniczając się tylko do promieni bliskich osi. Dla przejrzystości rysunku kreślimy bieg promieni w soczewce stosunkowo grubej (rys. 24.30), pamiętać jednak musimy o tym zasadniczym założeniu, że grubość soczewki jest mała w porównaniu z wartościami promieni krzywizny r_1 i r_2.

Punkty O_1 i O_2 oznaczają środki krzywizn obu powierzchni soczewki. Prosta łącząca O_1 i O_2 przechodzi przez środek soczewki i stanowi oś główną soczewki. Na osi

głównej znajduje się punkt świecący A. Prześledzimy bieg jednego z promieni, idącego z punktu A i załamującego się w soczewce. Promień ten (AM) tworzy z osią główną mały kąt α. W punkcie padania promienia na soczewkę prowadzimy płaszczyznę styczną i normalną MO_1. Przechodząc do ośrodka optycznie gęstszego, promień załamuje się do normalnej i pada w punkcie N na drugą powierzchnię soczewki. W tym punkcie prowadzimy również płaszczyznę styczną i normalną NO_2. Normalna ta, podobnie jak i normalna MO_1, jest promieniem krzywizny odpowiedniej powierzchni soczewki. Promień załamany wychodzący z soczewki przecina oś główną w punkcie B. Do określenia położenia obrazu punktu A wykorzystujemy drugi promień, np. promień biegnący wzdłuż osi głównej. Promień ten pada na soczewkę prostopadle do jej powierzchni i przechodzi bez załamania. Oba promienie przecinają się po załamaniu w punkcie B tworząc tam obraz punktu A.

Rys. 24.30

Promień NB tworzy z osią główną mały kąt β. Z tego wynika, że i promienie krzywizn MO_1 i NO_2 tworzą z osią główną niewielkie kąty ε i γ. Płaszczyzny styczne do powierzchni soczewek, poprowadzone w punktach padania i wyjścia rozpatrywanego promienia, tworzą ściany boczne zastępczego pryzmatu o kącie łamiącym φ. Ten sam kąt tworzą normalne do obu ścian:

$$\sphericalangle O_1CN = \varphi.$$

Jest to równocześnie kąt zewnętrzny w trójkącie O_2CO_1. A zatem

$$\varphi = \gamma+\varepsilon.$$

Przedłużenia promienia wchodzącego do pryzmatu i wychodzącego z pryzmatu tworzą kąt KEN równy kątowi odchylenia δ. Jest on kątem zewnętrznym w trójkącie AEB, czyli

$$\delta = \alpha+\beta.$$

Dzięki założeniu, że kąt łamiący φ i kąt padania promienia AM są małe, można zastosować wzór

$$\delta = (n-1)\varphi.$$

Otrzymujemy

$$\alpha+\beta = (n-1)(\gamma+\varepsilon). \tag{24.9}$$

Wobec tego, że kąty α, β, γ i ε są małe, możemy je zastąpić ich sinusami lub tangensami. Z założenia, że soczewka jest cienka, wynika, że wysokości punktów M i N nad osią główną są sobie w przybliżeniu równe:

$$MD \approx NG = h.$$

A zatem oznaczając odległość przedmiotu od soczewki przez x, a odległość obrazu od soczewki przez y, otrzymamy

$$\alpha \approx \operatorname{tg}\alpha = \frac{h}{x}, \qquad \varepsilon \approx \sin\varepsilon = \frac{h}{r_1},$$

$$\beta \approx \operatorname{tg}\beta = \frac{h}{y}, \qquad \gamma \approx \sin\gamma = \frac{h}{r_2}.$$

Podstawiając te wartości do wzoru (24.9) otrzymamy

$$\frac{h}{x}+\frac{h}{y} = (n-1)\left(\frac{h}{r_1}+\frac{h}{r_2}\right),$$
$$\frac{1}{x}+\frac{1}{y} = (n-1)\left(\frac{1}{r_1}+\frac{1}{r_2}\right). \tag{24.10}$$

Lewa strona równania ma tę samą postać co we wzorze dla zwierciadeł kulistych. W celu zachowania analogii wprowadzamy oznaczenie

$$(n-1)\left(\frac{1}{r_1}+\frac{1}{r_2}\right) = \frac{1}{f}. \tag{24.11}$$

Wielkość f określoną tym równaniem nazywamy *ogniskową soczewki.* Jest to odległość od soczewki takiego punktu, w którym przecinają się po załamaniu promienie tworzące wiązkę padającą, równoległą do osi głównej. Ogniskowa wyznacza położenie punktu F, zwanego *ogniskiem* (głównym) *soczewki.* W punkcie F przypada obraz źródła znajdującego się w nieskończoności.

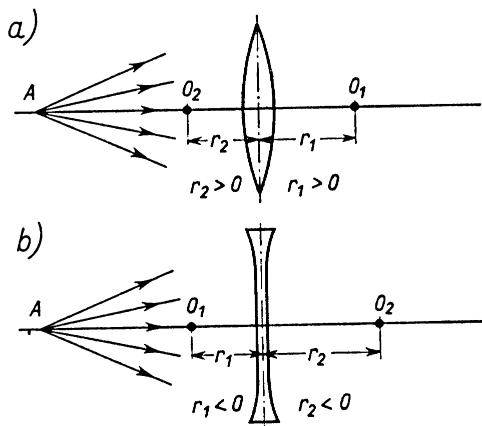
Jak widać z równania (24.11), długość ogniskowej zależy od współczynnika załamania materiału, z którego zrobiona jest soczewka, względem otaczającego ośrodka, jak również od promieni krzywizny obu jej powierzchni. Na przykład dla soczewki dwuwypukłej symetrycznej o promieniach krzywizn $r_1 = r_2 = r$ i współczynniku załamania $n = 1{,}5$ odległość ogniskowa równa się długości promienia r. Stąd wynika, że w odniesieniu do soczewek nie obowiązuje zależność $f = r/2$, z którą mieliśmy do czynienia w przypadku zwierciadeł sferycznych.

Omówimy obecnie zasadę, zgodnie z którą wielkościom r_1 i r_2 przypisujemy znaki plus lub minus.

Jeśli środek krzywizny O_1 znajduje się po przeciwnej stronie soczewki niż źródło światła A (rys. 24.31), to r_1 uważamy za dodatnie. Jeśli środek krzywizny O_2 znajduje się po tej samej stronie soczewki co źródło światła (przedmiot świecący), to r_2 uważamy

za dodatnie. Tak więc w przypadku soczewek dwuwypukłych oba promienie krzywizn są dodatnie. W soczewkach dwuwklęsłych oba promienie są ujemne. Dla powierzchni płaskich $r = \infty$.

Z wzoru na f wynika, że soczewki dwuwypukłe, sporządzone z materiału optycznie gęstszego niż otaczający ośrodek, mają ogniskowe dodatnie. Jest to charakterystyczne dla soczewek skupiających. Gdyby jednak soczewka była sporządzona z materiału optycznie rzadszego niż otaczający ośrodek, wtedy współczynnik załamania byłby mniejszy od jedności i ogniskowa takiej soczewki dwuwypukłej byłaby ujemna. Takie soczewki działają rozpraszająco.



Rys. 24.31

24.13. Dyskusja wzoru soczewkowego

Wprowadzając pojęcie ogniskowej soczewki otrzymaliśmy wzór dla soczewek cienkich w postaci

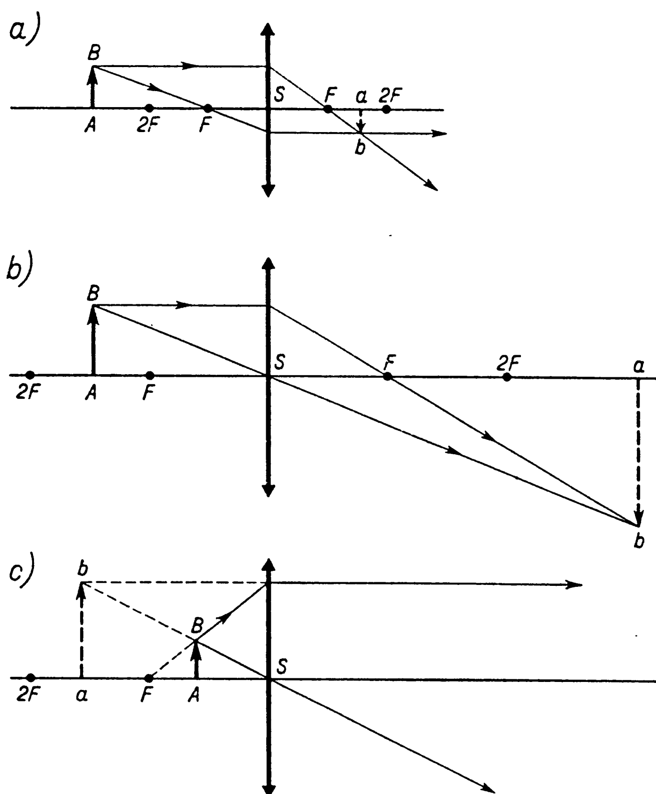
$$\frac{1}{x}+\frac{1}{y}=\frac{1}{f}.$$

Wobec identycznej postaci tego wzoru z wzorem dla zwierciadeł kulistych otrzymujemy te same wyniki dyskusji, z tym że wnioski dotyczące zwierciadeł skupiających, czyli wklęsłych, o ogniskowych większych od zera, są słuszne dla soczewek skupiających, czyli wypukłych. Natomiast wnioski dotyczące zwierciadeł rozpraszających, czyli wypukłych ($f < 0$), są słuszne dla soczewek rozpraszających, czyli wklęsłych. Trzeba jednak pamiętać, że nie obowiązuje zależność $f = r/2$.

W odniesieniu do soczewek należy kolejno rozważać następujące odległości x: $x > 2f$, $x = 2f$, x zawarte między f i $2f$, $x = f$, $x < f$. W związku z tym wygodnie jest uwzględniać na osi głównej punkt dodatkowy „$2F$", leżący w odległości równej $2f$ od soczewki. Odległość tę nazywać będziemy podwójną ogniskową.

Położenie punktu świecącego względem ogniska lub punktu „$2F$" decyduje o charakterze otrzymywanego obrazu.

W celu graficznego przedstawienia obrazów, tworzących się w soczewkach cienkich, należy schematycznie zaznaczyć soczewkę, nakreślić jej oś główną i zaznaczyć po obu stronach soczewki punkty F i $2F$. Przedmiot świecący, w postaci np. strzałki AB prostopadłej do osi głównej, umieszczamy poza podwójną ogniskową (rys. 24.32a). Obraz punktu B znajdujemy biorąc pod uwagę bieg dwóch w zasadzie dowolnych promieni. Wybieramy je tak, aby znany był ich bieg po załamaniu w soczewce. Wiadomo na przykład, że promień biegnący równolegle do osi głównej po załamaniu w soczewce wypukłej przechodzi przez ognisko rzeczywiste (leżące po przeciwnej stronie soczewki niż źródło światła). Taki sam promień po załamaniu w soczewce wklęsłej wychodzi z soczewki w takim kierunku, że jego przedłużenie przechodzi przez ognisko pozorne, leżące po tej samej stronie soczewki co przedmiot świecący (rys. 24.33a).

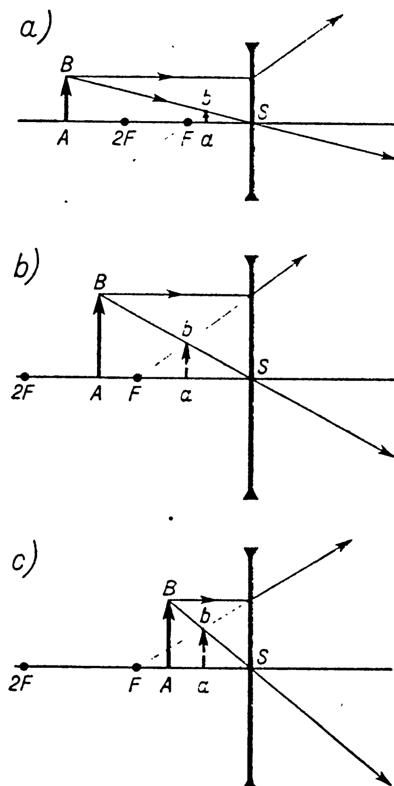


Rys. 24.32

Promień padający przechodzący przez ognisko soczewki wypukłej wychodzi po załamaniu jako promień równoległy do osi głównej.

Przy kreśleniu obrazów w soczewkach bardzo pomocny jest punkt S, zwany *środkiem optycznym* soczewki cienkiej. Jest to taki punkt, przez który promień przechodzi bez załamania. Wykorzystaliśmy ten punkt kreśląc rys. 24.32b i c oraz 24.33a, b i c.

Rysunki 24.32a, b, c przedstawiają bieg promieni w soczewkach zbierających (wypukłych), a rysunki 24.33a, b i c odnoszą się do soczewek rozpraszających (wklęsłych). Na tych ostatnich rysunkach widzimy, że zmiana odległości źródła od soczewki rozpraszającej nie wpływa na charakter obrazu (zawsze obraz jest pozorny), decyduje jednak o jego rozmiarach poprzecznych.



Rys. 24.33

Z podobieństwa trójkątów *ABS* i *Sab* (rys. 24.32b) wynika bezpośrednio wzór na powiększenie liniowe soczewki:

$$W = \frac{ab}{AB} = \frac{y}{x}.$$

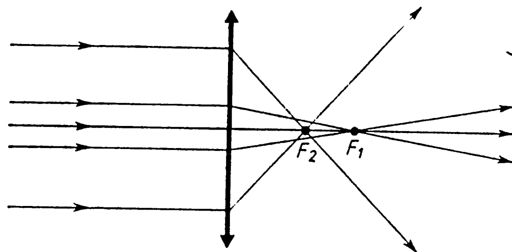
Im dalej od soczewki tworzy się obraz, tym jego rozmiary poprzeczne są większe.

24.14. Wady odwzorowań w soczewkach

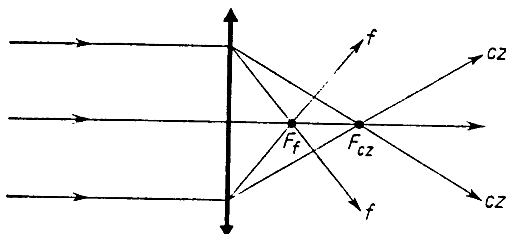
Podstawowymi wadami odwzorowań optycznych w soczewkach są: aberracja sferyczna, aberracja chromatyczna i astygmatyzm. Omówimy kolejno ich istotę.

Jeśli na soczewkę pada szeroka wiązka promieni równoległych, to po załamaniu

w soczewce promienie nie przecinają się w jednym punkcie, nie ma jednego punktowego ogniska dla całej szerokiej wiązki. Promienie biegnące blisko osi głównej (rys. 24.34) przecinają się dalej od soczewki w ognisku F_1, promienie skrajne przecinają się bliżej soczewki w ognisku F_2. Odległość F_1F_2 przyjęto za miarę *aberracji sferycznej podłużnej*. Jeśli w ognisku F_1 umieścimy ekran prostopadle do osi optycznej, to zamiast obrazu punktowego otrzymamy obraz w postaci krążka. Promień tego krążka jest miarą *aberracji sferycznej poprzecznej*.



Rys. 24.34



Rys. 24.35

Aberrację sferyczną można zmniejszać stosując zamiast soczewek pojedynczych układy wielosoczewkowe o odpowiednio dobranych promieniach krzywizny. Ustawienie soczewek o budowie niesymetrycznej ma również wpływ na wielkość aberracji sferycznej. Mniejszą aberrację sferyczną wykazuje soczewka wtedy, gdy załamanie światła rozkłada się mniej więcej równomiernie na obie powierzchnie soczewki. Na przykład ustawiając soczewkę płasko-wypukłą stroną płaską względem padającej wiązki równoległej otrzymujemy większe zniekształcenie obrazu wskutek aberracji sferycznej niż ustawiając tę samą soczewkę stroną wypukłą. Układy soczewek wolne od aberracji sferycznej nazywamy *układami aplanatycznymi*.

Drugą podstawową wadą odwzorowań w soczewkach jest *aberracja chromatyczna* związana z rozszczepieniem światła przy załamaniu, czyli wynikająca z faktu, że współczynnik załamania n jest funkcją częstotliwości (barwy) padającego promieniowania. Weźmy pod uwagę dwa skrajne promienie światła białego (rys. 24.35) padające na soczewkę równolegle do osi głównej. Po załamaniu (podobnie jak w pryzmacie) większe

odchylenie od pierwotnego biegu wykaże składnik fioletowy światła białego, mniejsze — składnik czerwony. A zatem w różnych punktach na osi głównej przypadną ogniska F_f i F_{cz}, przy czym ognisko F_{cz} będzie leżało dalej od soczewki. Odległość F_fF_{cz} jest miarą *aberracji chromatycznej podłużnej*.

Ustawiając ekran prostopadle do osi głównej w punktach F_f lub F_{cz} otrzymamy obraz źródła z barwną obwódką czerwoną lub fioletową. Przesuwając ekran od jednego ogniska do drugiego będziemy mieli obraz w postaci barwnego krążka, zmieniającego rozmiary w miarę przesuwania ekranu. Najmniejszą średnicę tego barwnego krążka przyjmujemy za miarę *aberracji chromatycznej poprzecznej*.

W celu zmniejszenia aberracji chromatycznej stosuje się również układy soczewek zamiast soczewek pojedynczych. Odpowiednie zestawienie soczewek zbierających i rozpraszających sporządzonych z różnych gatunków szkła może dać całkowitą *achromatyzację* dla pewnych długości fal (por. § 28.3).

Usunięcie aberracji chromatycznej ma duże znaczenie w optycznych przyrządach soczewkowych. W zależności od tego, czy przyrządy te są przeznaczone do badań wyłącznie wizualnych, wyłącznie fotograficznych czy też mieszanych, stosujemy układy achromatyczne dla różnie dobranych długości fal. W przyrządach przeznaczonych do obserwacji wizualnej, np. w mikroskopie, tak dobieramy układ soczewek, aby był achromatyczny dla określonej długości fali z dziedziny czerwonej i niebieskiej widma. Okazuje się, że taki układ odznacza się stosunkowo małą aberracją chromatyczną dla tych długości fal, na które oko ludzkie jest najbardziej czułe, tzn. dla barw widma między czerwoną i niebieską.

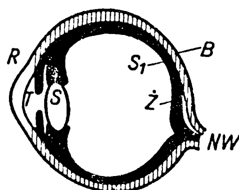
Wobec maksymalnej czułości kliszy na promienie fioletowe i niebieskie stosujemy achromatyzację dla określonych długości fal z tych obszarów widmowych w takich przyrządach, w których stosujemy wyłącznie fotograficzną metodę badań (np. refraktory astronomiczne). Jeśli badania są prowadzone częściowo wizualnie, a częściowo fotograficznie, to dążymy do achromatyzacji dla określonych długości fal w dziedzinie żółtej i fioletowej widma. Wtedy wady wynikające z aberracji chromatycznej są stosunkowo niewielkie dla zakresu dużej czułości oka i kliszy.

Warto tu wspomnieć, że achromatyzację można też osiągnąć przez rozsunięcie soczewek zbierających. Dwie soczewki zbierające o ogniskowych f_1 i f_2, umieszczone w odległości $(f_1+f_2)/2$, stanowią układ achromatyczny.

Oprócz takich wad odwzorowań w soczewkach, jak aberracja sferyczna i chromatyczna, istnieją jeszcze inne wady. Na przykład obrazem punktu świecącego, leżącego daleko od osi głównej, obejmującego swym promieniowaniem dużą powierzchnię soczewki, jest figura przypominająca kształtem przecinek. Jest to tzw. *wada koma*. Wiązka promieniowania wychodząca z punktu świecącego odległego od osi głównej, a więc padająca mocno skośnie na środkową część soczewki, również nie daje punktowego obrazu. Powstaje obraz w postaci dwóch wichrowatych, wzajemnie do siebie prostopadłych odcinków. Tę wadę odwzorowań optycznych nazywamy *astygmatyzmem*. Astygmatyzm występuje w przypadkach, gdy soczewka nie wykazuje symetrii względem osi padającej wiązki. Soczewki pochylone względem osi padających wiązek, jak również soczewki walcowe, wykazują zawsze odkształcenia astygmatyczne powstających obrazów.

24.15. Przyrządy optyczne

24.15.1. *Oko.* Przed przystąpieniem do omówienia budowy i zasady działania kilku przyrządów optycznych, skonstruowanych przez człowieka, parę uwag poświęcimy budowie oka ludzkiego i procesowi widzenia.



Rys. 24.36

Na rysunku 24.36 przedstawiono schematycznie budowę oka ludzkiego. Widzimy tam część zewnętrzną, tzw. białkówkę *B*, zwaną też twardówką. Przednia, przezroczysta część białkówki stanowi rogówkę *R*. Poza rogówką od strony wnętrza oka znajduje się tęczówka *T* stanowiąca przednią część warstwy zwanej naczyniówką. Przestrzeń między rogówką i tęczówką wypełniona jest płynem wodnistym. U różnych ludzi tęczówka jest różnie zabarwiona. W tęczówce znajduje się otwór zwany źrenicą, przez który promienie świetlne docierają do soczewki ocznej *S* znajdującej się głębiej. Poza soczewką znajduje się ciało szkliste o współczynniku załamania około 1,336, a więc prawie równym współczynnikowi załamania wody. Za ciałem szklistym rozciąga się czuła na światło warstwa zwana siatkówką S_1. Promienie świetlne przechodzą przez złożony układ optyczny składający się z rogówki, płynu wodnistego, soczewki i ciała szklistego. Po załamaniu promienie wytwarzają na siatkówce obraz rzeczywisty, pomniejszony i odwrócony. Na powierzchni siatkówki wyróżniamy dwa charakterystyczne miejsca, zwane plamką ślepą i plamką żółtą. Plamka ślepa jest niewrażliwa na światło i przypada w pobliżu miejsca wejścia nerwu wzrokowego *NW* do oka. Plamka żółta *Ż* jest najbardziej wrażliwa na światło. Reakcja oka na barwy i na mniejsze lub większe natężenie światła wiąże się z istnieniem na siatkówce zakończeń nerwu wzrokowego w postaci czopków i pręcików. Istnieje przypuszczenie, że zmianom zachodzącym w czopkach pod wpływem światła zawdzięczamy wrażenia barwne.

Oko ludzkie reaguje na światło w bardzo szerokich granicach natężeń — w stosunku około 1: 100 000. Najmniejsza ilość (dawka) energii wywołująca wrażenie światła, wynosi $2 \cdot 10^{-17}$ dżuli, co odpowiada około 30 fotonom światła żółtozielonego.

Na bardzo silne wiązki świetlne oko reaguje automatycznie skurczem mięśnia, zmniejszającym rozmiary źrenicy. W przypadku małych natężeń światła źrenica się rozszerza. Tę zdolność przystosowywania się oka do odbioru silniejszych i słabszych wiązek świetlnych nazywamy *adaptacją* oka.

Oprócz zdolności adaptacji oko (szczególnie w młodym wieku) ma zdolność *akomodacji*, tzn. zdolność dostosowywania się do odległości oglądanych przedmiotów. Człowiek o normalnym wzroku widzi ostro przedmioty zarówno stosunkowo bliskie (od-

ległość około 20 cm), jak i dalekie, ponieważ obrazy tych przedmiotów tworzą się na siatkówce. Jest to możliwe dzięki temu, że oko dostosowuje się do odległości oglądanych przedmiotów przez odpowiednią zmianę promieni krzywizny soczewki ocznej, wywołaną działaniem mięśnia ocznego. Promienie krzywizny powierzchni soczewki ocznej mogą się wahać w granicach od 5,7 do 10,7 mm.

Punkt najdalszy oglądany bez działania mięśnia akomodującego nosi nazwę *punktu dalekiego.* Dla oka normalnego przypada on w nieskończoności. Przedmioty bliższe wymagają już lekkiej akomodacji oka, ale jeszcze przy odległości 25 cm człowiek o wzroku normalnym nie odczuwa napięcia mięśnia. Odległość 25 cm przyjmowana jest za średnią odległość wyraźnego widzenia. Oglądanie przedmiotów z odległości mniejszej od 15 cm wymaga już napięcia mięśnia. Punkt taki, który jeszcze wyraźnie widzimy przy najsilniejszym napięciu mięśnia akomodującego, nosi nazwę *punktu bliskiego.* Jego odległość dla normalnego oka wynosi około 10 cm.

Na zakończenie krótka wzmianka o najbardziej rozpowszechnionych wadach wzroku, którymi są: dalekowzroczność i krótkowzroczność. U dalekowidzów przy całkowitym rozluźnieniu mięśnia ocznego punkt zbieżności wiązki promieni równoległych przypada poza siatkówką, a nie tak jak u ludzi o wzroku normalnym — na samej siatkówce. Korygowanie wady dalekowidztwa polega na użyciu okularów o szkłach wypukłych. Po przejściu przez soczewkę okularów wiązka promieni zmniejsza swą rozbieżność, dzięki czemu po dodatkowym załamaniu w soczewce ocznej punkt zbieżności wiązki zostaje przesunięty na siatkówkę.

U krótkowidzów przy całkowitym odprężeniu mięśnia punkt zbieżności wiązki promieni równoległych przypada przed siatkówką. W tym przypadku trzeba przez użycie okularów wklęsłych powiększyć rozbieżność wiązki, aby przesunąć punkt zbieżności na siatkówkę.

Wadę astygmatyzmu soczewki ocznej koryguje się przez zastosowanie soczewek walcowych.

Gdy mowa jest o korekcji wzroku, warto przypomnieć pojęcie zdolności zbierającej soczewek i jednostkę zdolności zbierającej, a mianowicie *dioptrię. Zdolność zbierająca* soczewki wyrażona w dioptriach jest odwrotnością odległości ogniskowej podanej w metrach. Na przykład: +5 dioptrii ma soczewka skupiająca o ogniskowej 20 cm, a −5 dioptrii ma soczewka rozpraszająca o ogniskowej −20 cm. Zestaw soczewek o ogniskowych f_1, f_2, f_3 itd. ma zdolność zbierającą $1/f$ równą sumie zdolności zbierających soczewek wchodzących w skład zestawu, czyli

$$\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2} + \frac{1}{f_3} \dots, \tag{24.12}$$

gdzie f oznacza ogniskową układu.

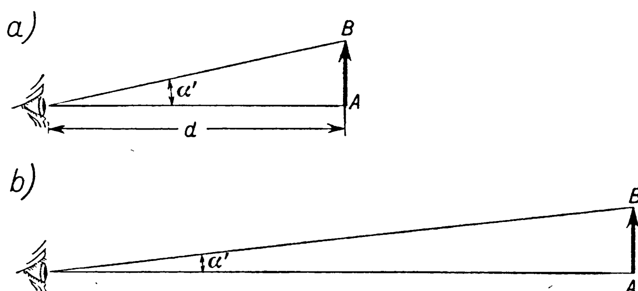
24.15.2. *Lupa.* Najprostszym przyrządem soczewkowym jest lupa, dająca obraz powiększony, pozorny i prosty. Lupa składa się z pojedynczej soczewki lub układu soczewek działających jak soczewka zbierająca.

Posługiwanie się lupą ułatwia człowiekowi odróżnianie szczegółów drobnych przedmiotów. Chcąc bez lupy dokładnie widzieć drobny przedmiot przybliżamy go do oka.

Wtedy wzrasta kąt, jaki tworzą graniczne promienie wpadające do oka (tzw. *kąt widzenia*), i wzrasta obraz oglądanego przedmiotu wytwarzany na siatkówce (rys. 24.37a i b). Wraz ze wzrostem kąta widzenia α' obraz dowolnego przedmiotu staje się bardziej odległy od obrazu punktu sąsiedniego. W związku z tym dostrzegamy więcej szczegółów. Ograniczona zdolność akomodacji decyduje jednak o tym, że kąta widzenia nie można dowolnie zwiększać przez zbliżanie przedmiotu do oka, gdyż po przekroczeniu pewnej odległości obraz przestaje być ostry.

Szczegółów przedmiotu nie dostrzegamy, jeżeli obraz na żółtej plamce nie pobudza na raz kilku czopków. Odległość czopków w sąsiedztwie żółtej plamki oceniana jest na 0,004 mm. Obraz tej wielkości tworzy się wtedy, gdy oglądany przedmiot widzimy pod kątem równym 1 minucie kątowej, czyli około $\frac{1}{3600}$ radiana. Jest to więc najmniejszy kąt widzenia, przy którym jeszcze ostro widzimy szczegóły przedmiotu.

W celu zwiększenia kąta widzenia korzystamy pomocniczo z takich przyrządów optycznych jak lupa i mikroskop.



Rys. 24.37

Aby scharakteryzować zdolność lupy do zwiększania kąta widzenia, podajemy zwykle wartość tzw. *powiększenia kątowego*. Przez powiększenie kątowe W_k rozumiemy stosunek kąta α, pod którym widzimy dany przedmiot za pomocą przyrządu optycznego, do kąta α', pod którym widzimy ten sam przedmiot okiem nieuzbrojonym z odległości wyraźnego widzenia:

$$W_k = \frac{\alpha}{\alpha'}. \tag{24.13}$$

Powiększenie kątowe lupy można powiązać z jej odległością ogniskową. Rozpatrzymy w tym celu zależności geometryczne przedstawione na rys. 24.38.

Mały przedmiot AB znajduje się w takiej odległości x od lupy o ogniskowej f, że jego obraz pozorny ab tworzy się w odległości wyraźnego widzenia ($y = -d$). Załóżmy, że oko znajduje się bezpośrednio za lupą. Małe rozmiary przedmiotu AB i jego obrazu ab decydują o tym, że popełniamy niewielki błąd biorąc pod uwagę zamiast kątów widzenia α i α' ich tangensy, a zatem

$$\alpha \approx \frac{ab}{d}, \qquad \alpha' \approx \frac{aB'}{d} = \frac{AB}{d},$$

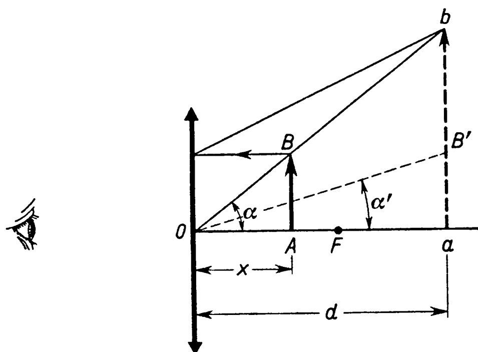
$$W_k = \frac{\alpha}{\alpha'} = \frac{ab}{AB}.$$

Z podobieństwa trójkątów Oab i OAB wynika proporcja

$$\frac{ab}{AB} = \frac{d}{x},$$

a więc

$$W_k = \frac{d}{x}. \tag{24.14}$$



Rys. 24.38

Uwzględniając ujemną wartość odległości obrazu pozornego od soczewki napiszemy

$$\frac{1}{x} - \frac{1}{d} = \frac{1}{f},$$

$$\frac{1}{x} = \frac{f+d}{df}.$$

Po podstawieniu do wzoru (24.14) otrzymujemy

$$W_k = \frac{d(d+f)}{df} = \frac{d}{f} + 1. \tag{24.15}$$

Innymi słowy, powiększenie kątowe lupy zależy od odległości wyraźnego widzenia obserwatora i odległości ogniskowej lupy. Lupy o krótszej ogniskowej dają większe powiększenia kątowe. Dalekowidz o większej odległości wyraźnego widzenia posługując się określoną lupą uzyskuje większe powiększenie kątowe niż krótkowidz.

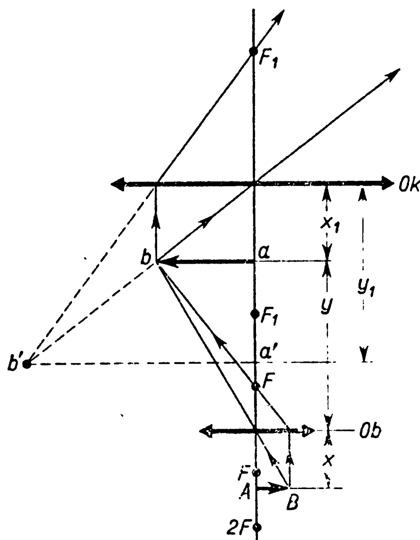
Pojedyncze soczewki zbierające użyte jako lupy dają powiększenie 8-10-krotne. Za pomocą układu soczewek, działającego jako lupa, można otrzymać powiększenia do 20 razy.

24.15.3. *Mikroskop*. Powiększenia uzyskane za pomocą lupy są niewystarczające, gdy oglądane przedmioty są bardzo małe. Posługujemy się wtedy mikroskopem. Schemat budowy mikroskopu przedstawiony jest na rys. 24.39.

W mikroskopie występują dwa układy soczewek: obiektyw (Ob) od strony ogląda-

nego przedmiotu i okular (Ok) od strony oka. Oba układy działają jak soczewki zbierające o odpowiednio zmniejszonej aberracji sferycznej i chromatycznej. Obiektyw ma bardzo krótką ogniskową, okular — nieco dłuższą.

Przedmiot AB umieszczamy między F i $2F$ obiektywu, bardzo blisko ogniska F, aby otrzymać duże powiększenie. Rzeczywisty obraz ab tworzy się w odległości większej od podwojonej odległości ogniskowej obiektywu. Okular ustawiamy w stosunku do tego obrazu jak lupę, tzn. w odległości nieco mniejszej od odległości ogniskowej okularu. Dzięki temu otrzymujemy obraz $a'b'$ pozorny, powiększony i prosty w stosunku do obrazu ab.



Rys. 24.39

Powiększenie liniowe W_l mikroskopu jest iloczynem powiększeń liniowych okularu i obiektywu:

$$W_l = W_{ob} W_{ok}. \tag{24.16}$$

Powiększenie kątowe W_k zdefiniowane jest wzorem analogicznym jak w przypadku lupy:

$$W_k = \frac{\alpha}{\alpha'}.$$

Wyrażając, tak jak poprzednio, oba kąty przez ich tangensy otrzymujemy

$$\alpha \approx \frac{a'b'}{d}, \qquad \alpha' \approx \frac{AB}{d},$$

zatem

$$W_k = \frac{a'b'}{AB}.$$

Z podobieństwa trójkątów na rys. 24.39 wynikają — jak to łatwo sprawdzić — następujące proporcje:

$$\frac{a'b'}{ab} = \frac{y_1}{x_1}, \qquad \frac{ab}{AB} = \frac{y}{x}.$$

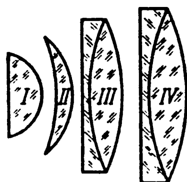
Przyjmując przybliżone równości

$$x_1 \approx f_{ok}, \quad y_1 \approx d, \quad x \approx f_{ob}, \quad y \approx L,$$

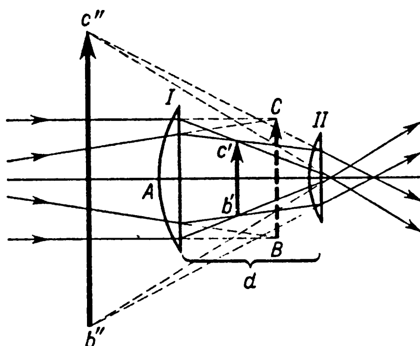
gdzie f_{ok} i f_{ob} oznaczają odpowiednio: ogniskową okularu i obiektywu, a L oznacza długość tubusa, otrzymujemy przybliżony wzór na powiększenie kątowe mikroskopu

$$W \approx \frac{dL}{f_{ob} f_{ok}}. \tag{24.17}$$

Z wzoru tego wynika, że mniejszym odległościom ogniskowym obiektywu i okularu odpowiada większe powiększenie kątowe mikroskopu. Graniczne wartości ogniskowych nowoczesnych mikroskopów wynoszą: dla obiektywów 1,5 mm, dla okularów 8,5–10 mm. Gdy długość tubusa wynosi 20 cm, a odległość wyraźnego widzenia 25 cm, otrzymujemy w tych warunkach powiększenie około 3000.



Rys. 24.40



Rys. 24.41

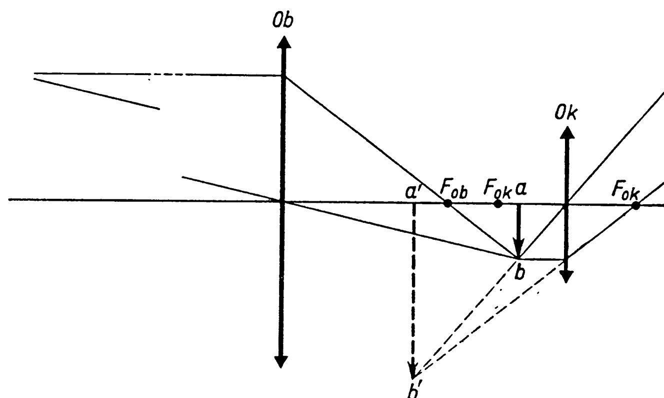
Obiektywy i okulary mikroskopów są układami soczewkowymi złożonymi, tak dobranymi, by zapewniały możliwie wysoką korekcję otrzymywanych obrazów. Nowoczesne precyzyjne obiektywy stanowią zestawy 8–10 soczewek, a obiektyw Amiciego (rys. 24.40), zbudowany w r. 1840, zapewniający duże powiększenie, składał się (już) z 6 soczewek. (Budowa pierwszego mikroskopu przypada prawdopodobnie na okres około r. 1590.)

Z okularów, które znalazły szerokie zastosowanie, wymienimy *okular Huygensa* (rys. 24.41). Składa się on z dwóch soczewek płasko-wypukłych o różnych ogniskowych f_1 i f_2 umieszczonych w odległości d i tak ustawionych, że wypukłe strony obu soczewek zwrócone są do źródła światła. Dzięki zachowaniu stosunku odległości

$$f_1 : d : f_2 = 3:2:1$$

ogniskowa całego układu wynosi $-\frac{3}{4}d$ (od A na prawo). Ognisko okularu leży zatem między soczewkami układu. Okular ten nie może służyć jako lupa do oglądania przedmiotów rzeczywistych. W mikroskopie umieszcza się go tak, aby obraz CB dawany przez obiektyw przypadał między soczewkami okularu. Obraz ten odgrywa wtedy rolę przedmiotu pozornego dla pierwszej soczewki okularu. Nie można jednak umieścić go na tle podziałki mikrometrycznej lub przecięcia nici pajęczych, gdyż w stosunku do nich tylko druga soczewka okularu spełniałaby rolę lupy.

W mikroskopie istotną rolę odgrywa odpowiednie oświetlenie przedmiotu. Na ogół nie wystarcza oświetlenie za pomocą ukośnego zwierciadła. Stosuje się zwykle kondensory rzucające na przedmiot szerokie wiązki zbieżne.

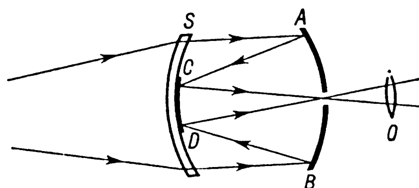


Rys. 24.42

Zastosowania mikroskopu w dzisiejszej nauce i technice są olbrzymie. Wymienimy przykładowo kilka dziedzin zastosowań, takich jak biologia, medycyna, metalurgia, chemia itp. Oprócz zwykłych mikroskopów optycznych budowane są różne mikroskopy specjalne, np. mikroskop interferencyjny, polaryzacyjny, mikroskop z kontrastem fazowym itp.

Do bardzo ważnego pojęcia zdolności rozdzielczej mikroskopu wrócimy jeszcze w rozdz. 25.

24.15.4. *Luneta astronomiczna.* Luneta astronomiczna, której działanie opiera się na zjawisku załamania światła, nosi nazwę refraktora. Schemat jej budowy przedstawiony jest na rys. 24.42. Składa się ona z 2 układów soczewek: obiektywu i okularu. Obiektyw ma dużą średnicę i długą ogniskową. Duże rozmiary obiektywu zapewniają udział licznych promieni w wytwarzaniu obrazu. Dzięki temu można otrzymywać obrazy nawet niezbyt jasnych ciał niebieskich. Promienie dochodzące do obiektywu tworzą bardzo mały kąt z osią optyczną wskutek bardzo dużej odległości obserwowanych ciał. Obraz *ab* odległego przedmiotu tworzy się między F i $2F$ obiektywu, praktycznie tuż za ogniskiem. Jest on zmniejszony, odwrócony i rzeczywisty. Oglądamy go przez drugą soczewkę — okular, umieszczony tak, aby ostateczny obraz *a'b'* był pozorny. Jest on zmniejszony w stosunku do oglądanego przedmiotu, ale znajduje się bliżej oka. Dzięki



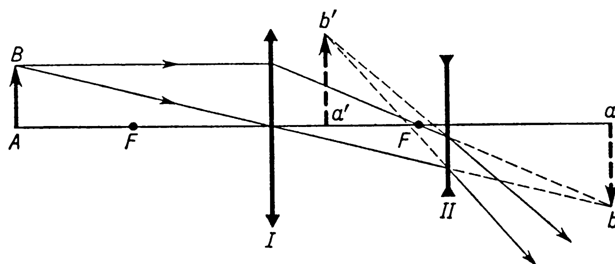
Rys. 24.43

temu luneta nie dając powiększenia liniowego, daje jednak powiększenie kątowe, umożliwiające dostrzeżenie większej liczby szczegółów na oglądanej powierzchni. Można udowodnić, że powiększenie kątowe lunety, określone jako stosunek kąta widzenia obrazu powstającego w lunecie do kąta widzenia przedmiotu, równa się stosunkowi odległości ogniskowych obiektywu i okularu:

$$W_k = \frac{f_{ob}}{f_{ok}}.$$

Obiektywy największych refraktorów astronomicznych mają średnicę około 1 m. Sporządzenie soczewki tak wielkich rozmiarów jest technicznie bardzo trudne. Dlatego też astronomowie stosują często lunety innego rodzaju, a mianowicie lunety, w których obiektywami są zwierciadła wklęsłe. Działanie tych reflektorów astronomicznych opiera się na prawach odbicia światła. Dodatkową korzyścią jest to, że w tym przypadku nie wchodzi w grę aberracja chromatyczna. Największy reflektor astronomiczny (teleskop zwierciadłowy) znajduje się w obserwatorium na Mount Palomar w Kalifornii. Zwierciadło tego reflektora ma średnicę 5 m.

Obecnie stosuje się również układy mieszane: soczewkowo-zwierciadlane. Schemat jednego z takich teleskopów przedstawiony jest na rys. 24.43. Widzimy tam soczewkę



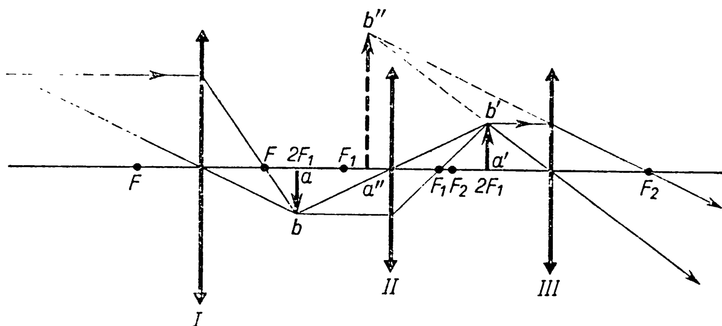
Rys. 24.44

wypukło-wklęsłą S i zwierciadło wklęsłe AB. Promienie po załamaniu w soczewce padają na zwierciadło AB i po odbiciu zostają skierowane na środkową część CD wklęsłej powierzchni soczewki S. Dzięki pokryciu warstwą aluminium ta część soczewki działa jak zwierciadło i odbija promienie po raz wtóry kierując je przez otwór w zwierciadle AB do okularu O. Układ daje obraz prosty, a więc może być z powodzeniem stosowany również przy obserwacjach na Ziemi.

24.15.5. *Luneta ziemska.* Lunety przeznaczone do obserwacji odległych przedmiotów na Ziemi muszą mieć inną budowę niż refraktory astronomiczne, gdyż nie mogą dawać ostatecznego obrazu odwróconego.

Lunety ziemskie są różnie konstruowane. Tak zwana luneta Galileusza składa się z obiektywu w postaci soczewki zbierającej *I* i okularu w postaci soczewki rozpraszającej *II* (rys. 24.44).

Promienie biegnące od odległych przedmiotów po załamaniu w obiektywie wytworzyłyby obraz *ab* rzeczywisty, odwrócony i zmniejszony, leżący prawie w płaszczyźnie ogniskowej obiektywu. (Na rysunku skala odległości nie może być zachowana.) Obraz taki jednak nie powstaje, gdyż na drodze tej wiązki zbieżnych promieni ustawiona jest soczewka rozpraszająca. Załamanie w tej soczewce zmienia wiązkę na rozbieżną, czyli powoduje powstanie obrazu *a'b'* pozornego i nieodwróconego w stosunku do przedmiotu. Kąt widzenia tego obrazu jest większy niż kąt widzenia odległego przedmiotu okiem nieuzbrojonym. Stosunek tych kątów jest miarą powiększenia kątowego lunety.



Rys. 24.45

Inny rodzaj lunety ziemskiej przedstawiony jest na rys. 24.45. Widzimy tam trzy soczewki zbierające. Soczewka *I* stanowi obiektyw i wytwarza obraz *ab* rzeczywisty, odwrócony i zmniejszony w stosunku do przedmiotu. Soczewka *II* odsunięta jest od tego obrazu o podwojoną odległość ogniskową. Dzięki temu jej działanie ogranicza się tylko do odwrócenia obrazu bez zmiany jego rozmiarów. Otrzymujemy nowy obraz rzeczywisty *a'b'*. Obraz ten jest przedmiotem dla soczewki *III* stanowiącej okular, umieszczonej w odległości mniejszej od ogniskowej od *a'b'*. Otrzymujemy zatem obraz *a''b''* prosty i pozorny. Ujemną stroną lunety tego typu jest jej duża długość.

Odwrócenie obrazu w lunecie ziemskiej można też uzyskać przez zastosowanie układu dwóch całkowicie odbijających pryzmatów, o czym była już mowa w § 24.9. Taką budowę mają lornetki pryzmatyczne.

Pytania i zadania

1. Jakie jest podstawowe pojęcie optyki geometrycznej? Jak można wyjaśnić to, że wyniki rozważań optyki geometrycznej są tylko przybliżone?
2. Jakie są podstawowe zjawiska analizowane w optyce geometrycznej? Jakie prawa nimi rządzą?

3. Jaka jest różnica między względnym i bezwzględnym współczynnikiem załamania? Jak można wyznaczyć względny współczynnik załamania dwóch ośrodków, jeżeli znamy ich bezwzględne współczynniki załamania?

4. Wyjaśnić tworzenie się obrazu w zwierciadle płaskim. Jakie są cechy utworzonego obrazu? Podać przykłady urządzeń, w których znajdują zastosowanie zwierciadła płaskie.

5. Podać szczegółowy plan wyprowadzenia wzoru wiążącego odległość przedmiotu i obrazu od zwierciadła z jego odległością ogniskową (definicja): a) dla promieni bliskich osi optycznej zwierciadła, b) dla promieni tworzących wiązki o skończonej rozwartości. Jaka wada odwzorowań w zwierciadłach ujawnia się w tym ostatnim przypadku?

6. Czy można otrzymywać obrazy w zwierciadłach bez aberracji sferycznej nawet w przypadku stosowania dość szerokich wiązek padających? Omówić dokładnie rodzaje zwierciadeł, które mogą służyć do tego celu.

7. Scharakteryzować wadę astygmatyzmu w zwierciadłach i soczewkach.

8. Wyjaśnić, na czym polega zjawisko całkowitego wewnętrznego odbicia? Jakie są warunki jego powstawania. Podać kilka zastosowań.

9. Wyjaśnić, dlaczego spoglądając na dno jeziora mamy złudzenie, że jest ono płytsze niż jest w rzeczywistości, dlaczego wiosło wydaje się złamane na granicy woda-powietrze, itp.

10. Przedstawić bieg promieni: a) w płytce płasko-równoległościennej, b) w pryzmacie, c) w cienkiej soczewce kulistej. Wyprowadzić dla każdego przypadku obowiązujące wzory.

11. Omówić tworzenie się obrazów w soczewkach skupiających i rozpraszających.

12. Co nazywamy zdolnością zbierającą soczewek? W jakich jednostkach wyraża się zdolność zbierająca?

13. Na czym polega i z czego wynika wada odwzorowań optycznych zwana aberracją chromatyczną? Czy wada ta występuje w zwierciadłach? Dlaczego?

14. Jakie są sposoby zmniejszania wad odwzorowań optycznych w soczewkach?

15. Omówić budowę oka ludzkiego i najpospolitsze wady wzroku. Jakie są sposoby pomagania ludziom cierpiącym na podstawowe wady wzroku?

16. Jaka jest budowa: a) lupy, b) mikroskopu, c) lunety astronomicznej i ziemskiej? Co nazywamy kątem widzenia i powiększeniem kątowym przyrządu optycznego? Od jakich parametrów zależy powiększenie kątowe lupy i mikroskopu?

17. Mały przedmiot umieszczono między dwoma płaskimi zwierciadłami zestawionymi tak, że kąt między ich płaszczyznami równa się 30°. Odległość przedmiotu od wspólnej krawędzi obu zwierciadeł $r = 10$ cm. Przedmiot znajduje się w dowolnym punkcie między zwierciadłami, lecz nie na dwusiecznej kąta. Obliczyć odległość dwóch pierwszych pozornych obrazów utworzonych w obu zwierciadłach.

Odp. 10 cm.

18. Promień światła pada na wklęsłe zwierciadło o promieniu R, równolegle do osi optycznej w odległości a od tej osi. Po odbiciu od zwierciadła przecina on oś optyczną w punkcie B. Obliczyć stosunek a/R, przy którym błąd względny wynikający z zaniedbania aberracji sferycznej wynosi 1%.

Odp. 0,14.

19. Na płytkę szklaną o współczynniku załamania 1,5 pada promień świetlny. Promień odbity od płytki tworzy z promieniem załamanym kąt 60°. Jaki jest kąt padania promienia na płytkę?

Odp. około 79°.

20. Spoglądając na dno zbiornika z wodą pionowo w dół odbieramy wrażenie, że jest on płytszy niż jest w rzeczywistości. Wykazać, że stosunek głębokości rzeczywistej do pozornej równa się współczynnikowi załamania.

21. Człowiek spogląda na swoje odbicie w zwierciadle płaskim umieszczonym na dnie zbiornika wypełnionego wodą na głębokości $h = 16$ cm (pod powierzchnią wody). Oko znajduje się na wysokości $H = 30$ cm nad powierzchnią wody. W jakiej odległości od oka znajduje się jego obraz pozorny wytworzony przez zwierciadło płaskie?

Odp. 84 cm.

22. Dwa ośrodki optyczne o współczynnikach załamania n_1 i n_2 graniczą ze sobą wzdłuż pewnej pła-

szczyzny. Jaką drogą powinien przejść promień świetlny z punktu A w ośrodku I do punktu B w ośrodku II, by czas przejścia był minimalny?

23. Na pryzmat o kącie łamiącym φ pada promień świetlny pod takim kątem α, że po załamaniu na przedniej ścianie trafia na ścianę tylną dokładnie pod kątem granicznym. Jaki jest współczynnik załamania szkła, z którego zrobiony jest pryzmat?

Odp. $n = \dfrac{\sqrt{1+2\sin\alpha\cos\varphi+\sin^2\alpha}}{\sin\varphi}$.

24. Zdolność zbierająca soczewki szklanej w powietrzu równa się 5,5 dioptrii. Jaka jest jej zdolność zbierająca po zanurzeniu do wody? Ile wynosi w tym drugim przypadku jej odległość ogniskowa?

Odp. 1,38 dioptrii; 0,72 m.

25. Na matówce aparatu fotograficznego otrzymano obraz o wysokości $h_1 = 60{,}6$ mm, gdy przedmiot znajdował się w odległości 15 m, a o wysokości $h_2 = 101{,}6$ mm, gdy przedmiot znajdował się w odległości 9 m. Jaka jest odległość ogniskowa obiektywu aparatu?

Odp. 13,2 cm.

26. Zdolność zbierająca lupy wynosi 10 dioptrii. W jakiej odległości od lupy powinien znajdować się przedmiot, aby obserwator, trzymający lupę przy oku, widział jego ostry obraz w odległości wyraźnego widzenia równej 25 cm?

Odp. 7,14 cm.

27. Z dwóch soczewek o ogniskowych $f_1 = 3$ cm i $f_2 = 4$ cm zestawiono model mikroskopu ustawiając je w odległości $d = 15$ cm. W jakiej odległości od pierwszej soczewki należy umieścić przedmiot, aby obserwator przykładając oko bezpośrednio do drugiej soczewki widział ostry obraz w odległości wyraźnego widzenia $d = 25$ cm?

Odp. 4,05 cm.

28. Obiektyw lunety Galileusza jest cienką soczewką dwuwypukłą o promieniach krzywizny $R_1 = R_2 = 24$ cm i o współczynniku załamania $n = 1{,}5$. Okularem jest soczewka dwuwklęsła o ogniskowej $f_2 = -5$ cm. Przedmiot znajduje się w odległości $x_1 = 30$ m od obiektywu. Jaka powinna być odległość obiektywu od okularu, aby obserwator o odległości wyraźnego widzenia $d = 25$ cm widział ostro ten przedmiot?

Odp. 17,95 cm.

Optyka falowa (I)
Interferencja i ugięcie światła

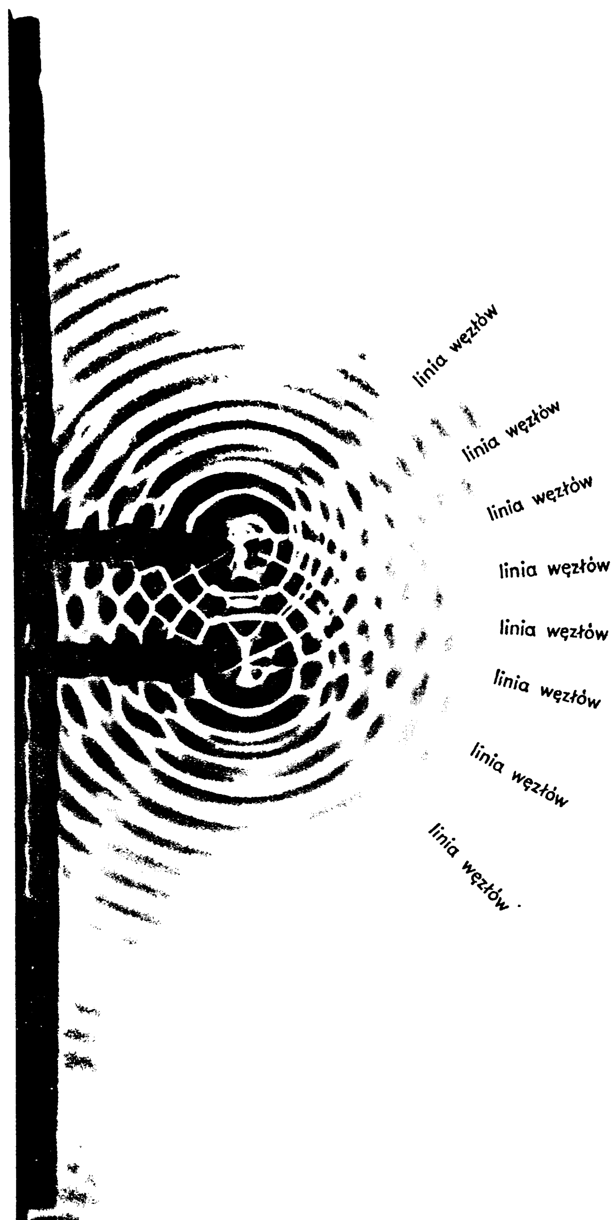
25.1. Rozważania wstępne

Jak już wspominaliśmy we wstępie do poprzedniego rozdziału, hipoteza falowej natury światła została wysunięta przez Huygensa mniej więcej w tym samym okresie, gdy Newton uzasadniał swą teorię korpuskularną. Obie teorie tłumaczyły takie zjawiska jak odbicie i załamanie. Słuszność wysuwanych przez obu uczonych przeciwstawnych wniosków dotyczących prędkości rozchodzenia się światła w różnych ośrodkach mogła być rozstrzygnięta dopiero po opracowaniu metod pomiaru prędkości światła w różnych ośrodkach, co jak wiemy, nastąpiło w wieku XIX (por. § 24.8). Gdyby w dziedzinie optyki udało się wykryć takie zjawiska, które można by wytłumaczyć stosując teorię falową, a nie można by wytłumaczyć na gruncie teorii korpuskularnej, sprawa byłaby przesądzona na rzecz teorii falowej. Okazało się, że takie zjawiska rzeczywiście istnieją. Można do nich zaliczyć interferencję, ugięcie (czyli dyfrakcję) i polaryzację światła. Są to zjawiska typowe dla ruchu falowego i właśnie one stanowią podstawy rozwoju teorii falowej światła.

25.2. Interferencja światła. Spójność promieniowania

Istotę zjawiska interferencji omawialiśmy dość szczegółowo w § 11.8. Ogólna jego charakterystyka pozostaje niezmieniona w odniesieniu do fal świetlnych. Natomiast trudności związane z obserwowaniem wyników interferencji światła wymagają dodatkowego omówienia.

Zjawisko interferencji światła występuje wtedy, gdy w określonym punkcie przestrzeni nakładają się dwie jednakowe, monochromatyczne fale świetlne (czyli fale o jednakowej częstotliwości). Fale te wzmacniają się lub osłabiają, zależnie od różnicy faz w miejscu spotkania. Najsilniejsze wzmocnienie towarzyszy nakładaniu się fal w fazach zgodnych, najsilniejsze osłabienie — nakładaniu się fal w fazach przeciwnych. Gdyby fale świetlne wychodziły z dwóch synchronicznie drgających źródeł, to każdemu punktowi ich spot-



Rys. 25.1

kania odpowiadałaby trwale utrzymująca się różnica faz zależna od różnicy dróg od obu źródeł do punktu spotkania. Można by wtedy, podobnie jak w przypadku interferujących fal na wodzie, wychodzących z 2 synchronicznie drgających źródeł (rys. 25.1), wyróżnić powierzchnie (linie) wygaszeń fal i wzmocnień. Byłyby to, jak łatwo sprawdzić, rodziny hiperbol w przypadku fal powierzchniowych na wodzie i rodziny hiper-

boloid w przypadku fal przestrzennych świetlnych. Obraz interferencji byłby statyczny i dzięki temu łatwo obserwowalny. Normalnie stosowane, niezależne od siebie źródła światła nie są jednak zsynchronizowane, w emisji światła każdego z nich występują nieregularne przeskoki fazowe, w związku z tym w każdym dowolnie wybranym punkcie ośrodka co chwila mogą się spotykać fale o coraz to innej różnicy faz, dające zupełnie inny wynik interferencji. Te zmiany fazy zachodzą z tak dużą częstotliwością, że człowiek nie zauważa szybko po sobie następujących wzmocnień i osłabień światła w danym punkcie, gdyż oko ludzkie ma pewną bezwładność: zachowuje wrażenie świetlne w ciągu mniej więcej dziesiątej części sekundy. Ta właściwość oka umożliwia widzenie np. w kinie ciągłości akcji zamiast oddzielnych obrazków, które w rzeczywistości są wyświetlane. Ta sama właściwość uniemożliwia zaobserwowanie kolejnych krótkotrwałych wzmocnień i osłabień wynikających z interferencji. Bezwładność przyrządów pomiarowych również utrudnia badanie wyników istniejącej interferencji. Gdyby jednak udało się w danym punkcie przestrzeni utrzymywać przez dostatecznie długi czas stałą różnicę faz nakładających się fal, to wynik interferencji można by zaobserwować. Gdyby różnica faz równała się całkowitej wielokrotności 2π, to wystąpiłoby najsilniejsze obserwowalne wzmocnienie światła. Gdyby zaś różnica faz odpowiadała nieparzystej wielokrotności π, wystąpiłoby najsilniejsze osłabienie światła, a nawet zupełne wygaszenie w przypadku jednakowych amplitud nakładających się fal.

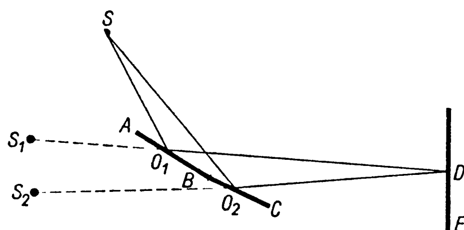
Doszliśmy więc do podstawowego *warunku zaobserwowania* (a nie warunku wystąpienia!) zjawiska interferencji w dziedzinie fal świetlnych. Interferencję światła obserwujemy wtedy, gdy w miejscu obserwacji utrzymuje się przez czas dostatecznie długi w porównaniu z okresem fali T stała różnica faz nakładających się fal. O takich falach mówimy, że są spójne.

Nasuwa się pytanie, kiedy spotykające się fale będą spełniały ten warunek? Trzeba tu przypomnieć, że fale świetlne są to fale elektromagnetyczne, polegające na rozchodzeniu się w przestrzeni pola elektrycznego i pola magnetycznego. W każdym punkcie przestrzeni, do którego dochodzi fala elektromagnetyczna, zachodzą okresowe zmiany wartości i kierunku wektora **E** natężenia pola elektrycznego i wektora **H** natężenia pola magnetycznego. Wektor **E** nazywamy czasem *wektorem świetlnym* lub *wektorem elektrycznym*, a wektor **H** — *wektorem magnetycznym*. W każdym punkcie przestrzeni, w której rozchodzi się fala elektromagnetyczna, oba wektory są wzajemnie do siebie prostopadłe i ulegają z biegiem czasu sinusoidalnym zmianom.

Każde ciało świecące, nawet tzw. źródło punktowe, składa się w rzeczywistości z wielkiej liczby źródeł. Wynika to stąd, że zmiany natężenia pola elektrycznego i magnetycznego związane są ze zmianami zachodzącymi w poszczególnych atomach. Światło wysyłane przez dowolne źródło jest to efekt sumaryczny zmian zachodzących w bardzo wielu atomach. W miejscu nakładania się fal wysyłanych przez dwa niezależne od siebie źródła światła nigdy nie utrzyma się stała różnica faz. Wynika to z nieregularności świecenia, z przypadkowych zmian faz zachodzących w obu źródłach. A zatem przy nakładaniu się fal pochodzących z dwóch niezależnych źródeł nie zaobserwujemy nigdy wyników interferencji w postaci wzmocnień i osłabień światła. O takich falach i o takich źródłach mówimy, że są *niespójne*, co właśnie oznacza, że przy nakładaniu

się fal rozchodzących się od takich źródeł nie otrzymamy trwale utrzymującej się stałej różnicy faz, czyli nie otrzymamy dającej się obserwować interferencji.

Z jednego źródła światła promienie rozchodzą się w rozmaitych kierunkach i na ogół już się nie przecinają. Stwarzając jednak na ich drodze warunki do wystąpienia np. zjawiska odbicia czy załamania możemy skierować promienie tak, by ich kierunki się przecięły. W miejscu nakładania się mamy fale spójne o trwale utrzymującej się różnicy faz, gdyż zawsze można brać pod uwagę dwa promienie rozchodzące się z tego samego atomu i dochodzące do punktu spotkania różnymi drogami (jak gdyby z dwóch pozornych źródeł światła) o stałej długości. Przy takim nakładaniu się fal wyniki interferencji można zaobserwować.



Rys. 25.2

Omówimy dokładniej sposób otrzymywania fal spójnych za pomocą odbicia od tzw. *zwierciadeł Fresnela*. Są to dwa zwierciadła płaskie AB i BC (rys. 25.2) zestawione w ten sposób, że tworzą kąt bliski 180°. Monochromatyczne źródło światła S ma obraz pozorny S_1 w zwierciadle AB i obraz pozorny S_2 w zwierciadle BC. Promienie SO_1 i SO_2, pochodzące ze źródła S po odbiciu od obu zwierciadeł, spotykają się w punkcie D na ekranie. Jeden z promieni biegnie jak gdyby ze źródła S_1, drugi — ze źródła S_2. Mimo że oba promienie wychodziły ze źródła S (z dowolnego atomu) w tej samej fazie, przy spotkaniu w punkcie D mogą one mieć fazy różne, gdyż przebyły różne drogi. Jeśli odległość S_2D różni się od odległości S_1D o całkowitą wielokrotność długości fali λ, to w punkcie D otrzymamy wzmocnienie światła (zgodność faz). Jeśli różnica dróg (S_2D-S_1D) równa się nieparzystej wielokrotności $\lambda/2$, otrzymamy wygaszenie światła: w punkcie D będzie ciemno. Danemu położeniu źródła S względem zwierciadła Fresnela odpowiada stała różnica dróg promieni dochodzących do punktu D i stała różnica faz nakładających się promieni. Wynik interferencji można zaobserwować.

Takie same rozważania można przeprowadzić dla innych punktów ekranu E. Jeśli źródło światła ma postać wąskiej szczeliny, to na ekranie zobaczymy na przemian ciemne i jasne prążki interferencyjne. Takie doświadczenie można zastosować do obliczenia długości fali świetlnej.

Stosując niemonochromatyczne źródła światła otrzymamy na ekranie prążki interferencyjne barwne. W przypadku różnych długości fal miejsca wzmocnienia promieniowania wskutek interferencji będą przypadały w różnych punktach ekranu.

W związku z opisanym doświadczeniem można jeszcze dodać, że przy odsuwaniu

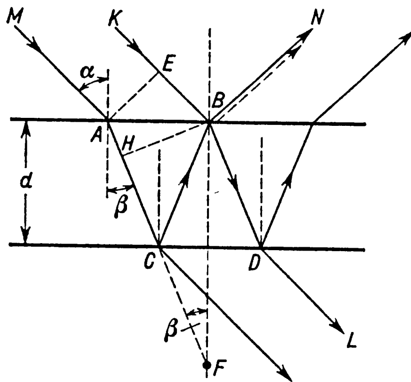
ekranu E od pozornych źródeł S_1 i S_2 w pewnym momencie prążki interferencyjne znikają, a pozostaje jednolicie oświetlona płaszczyzna ekranu. Wytłumaczenie tego zjawiska wiąże się z mechanizmem świecenia atomu. Zakłada się, że każdy atom w krótkim czasie swego pobudzenia wysyła ciąg fal o określonej częstotliwości. Średni „czas życia" atomów w stanie wzbudzonym jest rzędu 10^{-8} s. W tym czasie, jak łatwo obliczyć znając wartość okresu T fali świetlnej (rząd wielkości 10^{-15} s), zostaje wysłany ciąg składający się z około

$$\frac{10^{-8}}{10^{-15}} = 10^7 \text{ fal.}$$

Po czasie 10^{-8} s dany atom aż do nowego pobudzenia nie wysyła promieniowania. Nowy ciąg wysyłanych przez niego fal może mieć fazę zupełnie przypadkową w stosunku do fal pochodzących z poprzedniego pobudzenia. W przypadku nakładania się fal związanych z różnymi wzbudzeniami tego samego atomu nie będzie już spełniony warunek stałości faz w danym punkcie ekranu, a zatem fale nie będą spójne i interferencji nie będzie można zaobserwować. Ciąg 10^7 fal promieniowania np. zielonego o długości fali $5 \cdot 10^{-7}$ m rozciąga się na odległość $10^7 \cdot 5 \cdot 10^{-7}$ m $= 5$ m. A więc jeszcze przy różnicy odległości od punktu D ekranu do źródeł S_1 i S_2 wynoszącej 5 m można na nim obserwować prążki interferencyjne zielone. W przypadku dłuższych fal, jak również w przypadku atomów o dłuższym czasie życia, graniczna różnica długości, przy której jeszcze można obserwować obraz interferencyjny, jest większa.

25.3. Interferencja światła w cienkich warstwach

Przeanalizujemy bardziej szczegółowo przypadek interferencji światła występującej przy nałożeniu się na siebie promieni odbitych od dwóch powierzchni cienkiej płytki (błonki, warstwy cieczy itp.).



Rys. 25.3

Dla uproszczenia rozważań weźmy pod uwagę płytkę płasko-równoległą (rys. 25.3) o grubości d i o współczynniku załamania $n > 1$, umieszczoną w powietrzu. Na górną powierzchnię płytki pada fala płaska (biegnąca z punktu leżącego w nieskończoności),

przedstawiona w postaci dwóch równoległych promieni MA i KB. AE jest czołem tej fali w chwili dotarcia jej w punkcie A do powierzchni odbijającej. Rozważamy dalej: 1) promień BN (linia ciągła) powstający wskutek odbicia od ośrodka optycznie gęstszego, a więc o fazie zmienionej na przeciwną (por. § 11.10, zmiana fazy o π w przeliczeniu na drogę promienia daje $\lambda/2$) oraz 2) promień AC załamany w płytce pod kątem β, z którego dzięki odbiciu od ośrodka optycznie rzadszego (powietrza), a więc bez zmiany fazy powstaje promień CB, z którego z kolei wskutek załamania przy przejściu do powietrza powstaje promień BN (linia przerywana). Oba promienie biegnące wzdłuż BN są spójne, a więc zdolne do interferencji, której wynik zależy od różnicy dróg optycznych promieni $MACBN$ i KBN.

Drogi optyczne obu promieni liczone od źródła do czoła fali AE są jednakowe. Jednakowe są również drogi BN. Różnica Δ przebytych dróg optycznych sprowadza się zatem do wyrażenia

$$\Delta = (AC+BC)n-EB.$$

Ponieważ jednak HB jest nowym czołem fali załamanej, wytworzonym w chwili dotarcia fali do powierzchni płytki w punkcie B, więc drogi geometryczne EB i AH przebyte są w jednakowym czasie i stanowią jednakowe drogi optyczne, czyli

$$EB = nAH.$$

Zatem różnica dróg optycznych interferujących promieni biegnących wzdłuż BN sprowadza się do wyrażenia

$$\Delta = (HC+CB)n.$$

Stosując pomocniczą konstrukcję geometryczną, widoczną na rys. 25.3, możemy zastąpić

$$HC+CB = HC+CF = HF.$$

Ale z trójkąta HFB wynika, że

$$HF = 2d\cos\beta = 2d\sqrt{1-\sin^2\beta},$$

skąd po uwzględnieniu definicji współczynnika załamania otrzymujemy

$$HF = 2d\sqrt{1-\frac{\sin^2\alpha}{n^2}} = \frac{2d}{n}\sqrt{n^2-\sin^2\alpha},$$

a zatem

$$\Delta = 2d\sqrt{n^2-\sin^2\alpha}.$$

Uwzględniając zmianę fazy na przeciwną przy odbiciu od ośrodka optycznie gęstszego otrzymamy ostatecznie następujące wyrażenie na różnicę dróg interferujących promieni:

$$\delta=2d\sqrt{n^2-\sin^2\alpha} \pm \frac{\lambda}{2}.$$

Jak widać, różnica ta zależy od kąta padania α, od grubości płytki (błonki) i jej współczynnika załamania. Jeżeli różnica ta równa się całkowitej wielokrotności λ padającego

światła, to następuje wzmocnienie interferencyjne danej barwy. Jeżeli jednak różnica ta równa się nieparzystej wielokrotności $\lambda/2$, to następuje interferencyjne wygaszenie tej barwy. Warunki tworzenia się maksimów i minimów interferencyjnych w świetle odbitym od cienkich warstw (błonek) mają zatem następującą postać:

$$2d\sqrt{n^2-\sin^2\alpha} = (m+\tfrac{1}{2})\lambda, \text{ gdzie } m = 0, 1, 2, 3, \ldots \text{ (maksima)}, \quad (25.1)$$

$$2d\sqrt{n^2-\sin^2\alpha} = m\lambda, \text{ gdzie } m = 0, 1, 2, 3, \ldots \text{ (minima)}. \quad (25.2)$$

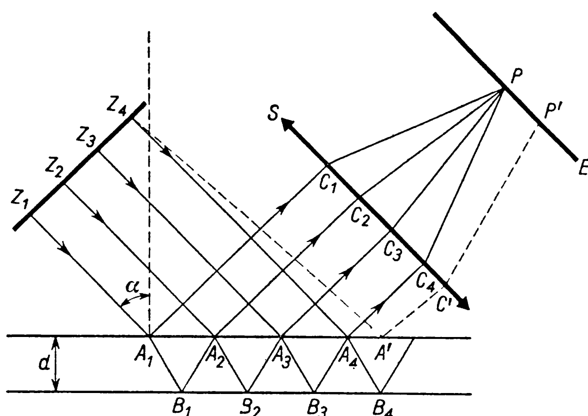
P y t a n i e. Jaką postać przyjmą powyższe wzory w przypadku prostopadłego padania światła na błonkę?

Jeśli grubość błonki nie jest wszędzie jednakowa, to do wzmocnienia interferencyjnego światła monochromatycznego, tzn. światła o określonej długości fali, przyczyniają się (przy stałym kącie padania α) tylko te punkty błonki, którym odpowiada ta sama grubość. Powstają wtedy linie wzmocnień (i osłabień) światła monochromatycznego, zwane *prążkami interferencyjnymi równej grubości*. Jeżeli taka błonka oświetlona jest światłem białym, a nie monochromatycznym, to wskutek różnych wzmocnień i osłabień interferencyjnych, zależnych od λ i od grubości błonki, zmieniają się stosunki natężeń składników światła białego w świetle dochodzącym do oka, co pociąga za sobą wrażenie barwne. Na przykład na błonce czy płytce w kształcie klina powstają prążki interferencyjne równej grubości w postaci linii prostych równoległych do krawędzi klina, ciemnych i jasnych w świetle monochromatycznym, a tęczowych — w świetle białym.

Ogólnie znane efekty baniek mydlanych, mieniące się różnobarwnie plamy smarów na mokrej jezdni, piękne barwy skrzydeł motyli — to właśnie przykłady barw interferencyjnych. Trzeba podkreślić, że efekty te mogą występować tylko wtedy, gdy błonki są naprawdę bardzo cienkie, tzn. gdy grubość d nie przekracza kilku długości fal świetlnych. Wyjaśnienie przyczyny tego zjawiska można znaleźć w obszerniejszych podręcznikach optyki.

Na zakończenie rozważań związanych z interferencją światła w cienkich błonkach zwrócimy jeszcze uwagę, że oprócz prążków równej grubości mogą jeszcze powstawać prążki interferencyjne, zwane prążkami równego nachylenia. Powstają one wtedy, gdy na płytkę płasko-równoległą o stałej grubości d pada światło pod wszystkimi możliwymi kątami. Źródło promieniowania nie jest tym razem źródłem punktowym w nieskończoności, lecz źródłem rozciągłym złożonym z wielkiej liczby elementów punktowych, umieszczonym w skończonej odległości od płytki. Weźmy pod uwagę kilka elementów punktowych Z_1, Z_2, Z_3 naszego źródła (rys. 25.4). Z elementów tych promienie padają pod różnymi kątami na powierzchnię cienkiej, płasko-równoległej płytki, weźmiemy jednak na razie pod uwagę tylko promienie padające na płytkę pod kątem α. Jak widać z rysunku, każdy z promieni Z_1A_1, Z_2A_2, Z_3A_3 częściowo ulega odbiciu na przedniej powierzchni płytki tworząc promienie A_1C_1, A_2C_2, A_3C_3, częściowo ulega załamaniu, wchodzi do płytki, odbija się od wewnętrznej dolnej powierzchni w punktach B_1, B_2, B_3, wraca do powierzchni górnej i znów częściowo się załamuje i wychodzi jako A_2C_2, A_3C_3 itd. lub kursuje po odbiciu między powierzchnią górną i dolną płytki itd. Można

zatem mówić o promieniach spójnych $Z_1A_1C_1$, $Z_1A_1B_1A_2C_2$, $Z_1A_1B_1A_2B_2A_3C_3$, które po skupieniu przez soczewkę S w punkcie P na ekranie E, umieszczonym w płaszczyźnie ogniskowej soczewki, dadzą pewien efekt interferencyjny (zależny od kąta α; $d = \text{const}$). Ale w tym samym punkcie P zostaną zebrane promienie biegnące z punktu Z_2, a więc spójne ze sobą, takie jak $Z_2A_2C_2$, $Z_2A_2B_2A_3C_3$ itd. Wynik ich interferencji (zależny znów tylko od kąta α) będzie taki sam, jak poprzedniej wiązki spójnych promieni. (Wiązka promieni pochodzących z Z_1 nie jest spójna z wiązką promieni



Rys. 25.4

idących z Z_2, Z_3 itd.) A zatem w punkcie P zostaną zebrane wszystkie promienie o tym samym kącie padania, o tym samym kącie nachylenia. Linią przerywaną zaznaczony jest na rysunku bieg jednego promienia $Z_4A'C'$ nachylonego pod innym kątem. Wiązka takich równoległych promieni byłaby już zebrana przez soczewkę S w innym punkcie P' płaszczyzny ogniskowej soczewki. A więc wiązki o jednakowym nachyleniu względem płytki wytworzą — po zebraniu ich w soczewce zbierającej (lub w nieskończoności) — prążki interferencyjne stałego nachylenia.

Zastosowanie praktyczne prążków stałej grubości i stałego nachylenia omówimy w § 25.5.

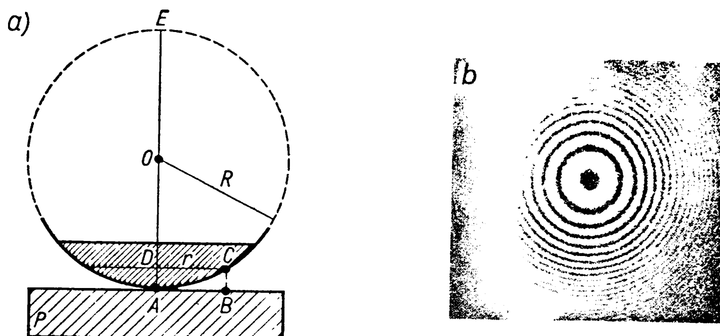
25.4. Pierścienie Newtona

Interesującego przykładu interferencji światła dostarcza doświadczenie z pierścieniami Newtona. Przez umieszczenie soczewki płasko-wypukłej o dużym promieniu krzywizny na płaskiej płytce szklanej uzyskujemy warstwę powietrza o stopniowo rosnącej grubości w miarę oddalania się od punktu A (rys. 25.5a). Monochromatyczne promienie równoległe, padające prostopadle na powierzchnię płaską soczewki, przechodzą przez szkło i częściowo ulegają odbiciu od powietrza, a częściowo przechodzą dalej przez warstwę powietrza, ulegają odbiciu od płytki dolnej P i dopiero wracają do oka obserwatora lub do obiektywu słabo powiększającego mikroskopu. Promienie

odbite np. w punkcie C i te, które dwukrotnie przeszły odległość $CB = d$ w powietrzu, interferują ze sobą. Różnica ich dróg optycznych wynosi

$$\delta = 2d + \frac{\lambda}{2}. \tag{25.3}$$

Występująca we wzorze wielkość $\lambda/2$ wynika z przeskoku fazy przy odbiciu od ośrodka optycznie gęstszego na powierzchni płytki P.



Rys. 25.5

Wobec symetrycznego układu warstwy powietrznej względem punktu A, obraz interferencyjny ma kształt koncentrycznych pierścieni (rys. 25.5b). Najbliższe otoczenie środka tych pierścieni zaznacza się na rysunku jako krążek ciemny. Jest to zgodne ze wzorem (25.3). Grubość warstewki powietrza w miejscu zetknięcia soczewki i płytki jest prawie równa zeru, a zatem różnica faz obu interferujących promieni sprowadza się do $\lambda/2$, co decyduje o wygaszeniu interferujących promieni.

Chcąc ustalić zależność między promieniami pierścieni jasnych i ciemnych i długością fali rozpatrujemy zależności geometryczne (rys. 25.5a). Oznaczmy promień rozważanego pierścienia $AB = r$, promień krzywizny soczewki $OA = R$, grubość warstwy powietrza nad punktem B, czyli $BC = d$.

Ze znanego twierdzenia geometrycznego wynika równość

$$DC^2 = ED \cdot DA, \quad \text{czyli} \quad r^2 = (2R-d)d.$$

Odrzucając w nawiasie d, jako bardzo małe w porównaniu z $2R$, otrzymujemy

$$r^2 = 2R\,d, \quad \text{skąd} \quad d = \frac{r^2}{2R}.$$

A zatem równanie (25.3) przyjmuje postać

$$\delta = \frac{r^2}{R} + \frac{\lambda}{2}.$$

Gdy ta różnica dróg optycznych równa się nieparzystej wielokrotności połówek długości fali, powstaje pierścień ciemny o promieniu r_k:

$$\frac{r_k^2}{R} + \frac{\lambda}{2} = (2k+1)\frac{\lambda}{2},$$

skąd

$$\frac{r_k^2}{R} = k\lambda, \quad \text{gdzie} \quad k = 1, 2, 3, \ldots$$

Promienie r_k ciemnych pierścieni spełniają zatem równanie

$$r_k = \sqrt{kR\lambda}, \quad \text{gdzie} \quad k = 1, 2, 3, \ldots$$

Warunek wytwarzania się pierścieni jasnych napiszemy w postaci

$$\frac{r_m^2}{R} + \frac{\lambda}{2} = m\lambda, \quad \text{gdzie} \quad m = 1, 2, 3, \ldots,$$

skąd

$$\frac{r_m^2}{R} = (2m-1)\frac{\lambda}{2},$$

czyli promienie pierścieni jasnych r_m spełniają równanie

$$r_m = \sqrt{R(2m-1)\frac{\lambda}{2}}, \quad \text{gdzie} \quad m = 1, 2, 3, \ldots$$

Promienie pierścieni jasnych, otrzymane przy użyciu określonej soczewki, są zatem proporcjonalne do pierwiastka kwadratowego z kolejnych liczb nieparzystych i do długości fali. Są zatem większe dla światła czerwonego niż dla fioletu. Przy użyciu światła białego do oświetlenia układu otrzymuje się układ paru pierścieni barwnych o brzegu wewnętrznym (o mniejszym r_m) fioletowym i brzegu zewnętrznym — czerwonym.

Doświadczenie z pierścieniami Newtona można wykorzystać do pomiaru długości fali. W celu zwiększenia dokładności pomiaru mierzy się zwykle (stosując mikroskop z okularem mikrometrycznym) promienie dwóch możliwie odległych pierścieni danej barwy, np. pierścienia n-tego i $(n+s)$-tego. Do obliczenia λ stosuje się wtedy — łatwy do wyprowadzenia — wzór

$$\lambda = \frac{r_{n+s}^2 - r_n^2}{sR}.$$

Warto podkreślić, że wyraźne barwne obrazy interferencyjne otrzymujemy tylko w przypadku *niewielkich grubości* warstw powietrza lub warstw przezroczystych, w których promienie, dzięki odbiciu „kursują” tam i z powrotem. Grubość takiej warstewki nie może przekraczać kilku długości fali.

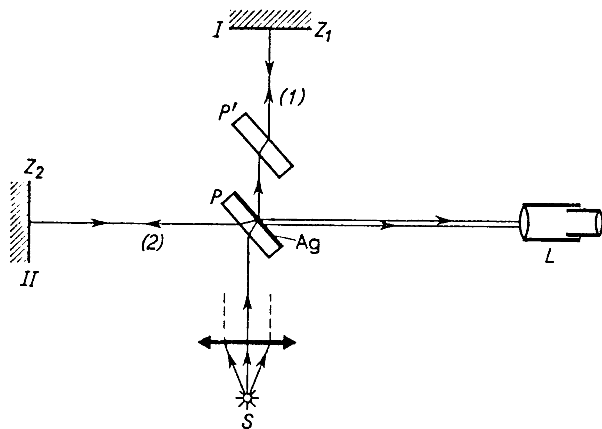
Pytanie. Czy układ pierścieni Newtona ulegnie zmianie, jeśli między soczewkę a płytkę ($n = 1{,}5$) wprowadzimy zamiast powietrza ciecz, np. olejek cedrowy lub goździkowy o współczynniku załamania $n = 1{,}6$?

25.5. Zastosowanie interferencji światła w metrologii

Zjawisko interferencji znalazło szerokie zastosowanie przede wszystkim w pomiarach bardzo małych odległości, jak również w pomiarach bardzo małych zmian odległości, rzędu nanometrów. Omówimy parę przykładów.

Jednym z ważniejszych przyrządów interferencyjnych, służących do pomiaru dłu-

gości fali i małych zmian długości, jest interferometr Michelsona (rys. 25.6). Równoległa wiązka światła monochromatycznego, biegnąca ze źródła *S*, pada na szklaną płasko-równoległą płytkę *P* pod kątem 45°. Jedna z powierzchni płytki pokryta jest półprzepuszczalną warstwą srebra. Dzięki temu na powierzchni tej następuje podział energii promienistej na dwie wiązki. Jedna z nich przechodzi przez warstwę srebra i biegnie dalej równolegle do pierwotnego kierunku, a druga — po odbiciu od warstwy srebra biegnie w lewo. Obie dochodzą do płaskich zwierciadeł *I* i *II* ustawionych prostopadle do kierunku dobiegających wiązek, toteż po odbiciu wracają one tymi samymi drogami do płytki szklanej, gdzie znowu dzięki warstwie półprzezroczystej każda z wiązek ulega częściowo odbiciu, a częściowo przechodzi przez płytkę. Dzięki temu z każdej wiązki część energii wraca w kierunku źródła, a druga część wędruje w stronę lunetki *L*. W lunetce obserwuje się wyniki interferencji dwóch nakładających się spójnych

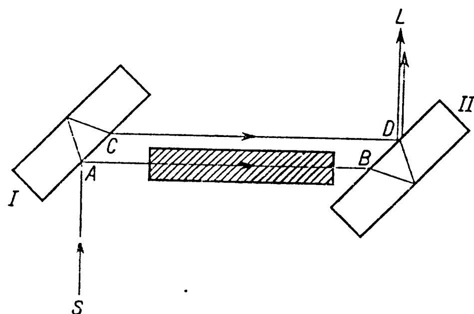


Rys. 25.6

wiązek. Prześledzenie biegu obu wiązek od chwili wyjścia ze źródła *S* do spotkania w lunetce wykazuje, że nawet przy równych odległościach zwierciadeł *I* i *II* od płytki *P* promień *2* przebywa drogę dłuższą, gdyż dwukrotnie „nadprogramowo" przechodzi przez szklaną płytkę *P*. Aby skompensować tę różnicę dróg optycznych obu promieni, ustawia się na drodze biegu promienia *1* równolegle do płytki *P* drugą płytkę *P'* identyczną z płytką *P*, tylko pozbawioną warstwy półprzezroczystej. Dzięki takiemu ustawieniu aparatury do lunetki docierają spójne wiązki dające w wyniku interferencji prążki interferencyjne równego nachylenia w postaci koncentrycznych pierścieni. Obraz interferencyjny jest taki sam, jaki powstałby w przypadku padania światła z rozciągłego źródła na zawartą między dwiema płytkami szklanymi warstwę powietrza o grubości *d*, równej różnicy odległości zwierciadeł *I* i *II* od płytki *P*. Jedno ze zwierciadeł może być przesuwane za pomocą precyzyjnej śruby mikrometrycznej. Łatwo można zdać sobie sprawę z tego, że przesunięcie tego zwierciadła o $\lambda/4$ zmienia drogę promienia o $\lambda/2$, czyli powoduje „przejście" w obrazie interferencyjnym od maksimum do minimum lub odwrotnie. Takie przejście np. w środku pola widzenia lunetki *L* może być

z łatwością odnotowane przez obserwatora, a zatem czułość przyrządu odpowiada $\frac{1}{4}$ długości fali, czyli przy użyciu światła fioletowego jest rzędu 10^{-7} m. Gdy odległości obu zwierciadeł od płytki P są jednakowe, w polu widzenia lunetki nie ma prążków interferencyjnych, pole widzenia jest jednolicie jasne. Stopniowa zmiana odległości powoduje pojawianie się prążków licząc od środka obrazu, przy czym łatwo sprawdzić, że przejście n prążków barwnych przez pole widzenia (co łatwo policzyć) odpowiada różnicy odległości zwierciadeł, równej $n\frac{\lambda}{2}$. Biorąc pod uwagę, że przesunięcia prążków można oceniać z dokładnością do $\frac{1}{100}$ ich grubości, można szacować dokładność pomiaru długości za pomocą interferometru na około 10^{-9} m.

Interferometr może służyć do porównywania długości badanych z długościami wzorcowymi, do pomiaru długości fali λ, do badania współczynników załamania gazów. Interferometr pozwolił Michelsonowi, a następnie innym badaczom na porównanie długości metra archiwalnego wzorcowego z długościami fal (por. definicję metra w układzie SI). Słynne doświadczenie Michelsona–Morleya, wykonane przy użyciu interferometru Michelsona, stwierdzające niezależność szybkości światła od ruchu Ziemi, obaliło hipotezę istnienia wszechświatowego eteru i stworzyło podstawy rozwoju teorii względności (por. rozdz. 7).

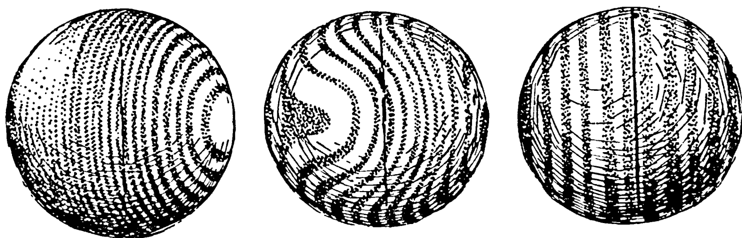


Rys. 25.7

Na badaniu interferencji opiera się też bardzo dokładny pomiar współczynnika rozszerzalności cieplnej. Tym razem przedmiot badany ustawia się tak względem powierzchni wzorcowej, aby między obu powierzchniami powstała cienka warstwa powietrza. W warstwie tej powstają prążki interferencyjne jednakowej grubości. Jeżeli badane ciało rozszerza się wskutek ogrzania, to grubość wspomnianej warstwy powietrza maleje i prążki zaczynają się przesuwać. Z przesunięcia prążków wnioskuje się o zmianie rozmiarów badanego ciała wyrażając ją w skali długości fali użytego światła. Metoda jest bardzo dokładna, gdyż przy użyciu światła monochromatycznego otrzymuje się prążki interferencyjne bardzo ostre i ich przesunięcia można mierzyć z dużą dokładnością (do 0,01 odległości między prążkami).

Specjalne urządzenie, zwane refraktometrem interferencyjnym (rys. 25.7), służy do dokładnego wyznaczania współczynnika załamania. Do otrzymania dwóch wiązek spójnych, rozdzielonych na pewnym odcinku ich drogi, służą dwie dość grube, jednako-

we, bardzo starannie oszlifowane płytki szklane, ustawione równolegle do siebie. Rysunek wyjaśnia mechanizm powstawania rozdzielonych wiązek na odcinkach AB i CD. Dzięki ich jednakowej długości (przy spełnieniu podanych wyżej warunków), nie ma żadnej różnicy dróg i oba promienie ulegają tylko wzmocnieniu podczas interferencji na końcu swej wspólnej drogi. Jeśli jednak ustawienie płytek nie jest idealnie równoległe, to występuje różnica dróg optycznych i w polu widzenia lunetki zbierającej L pojawiają się prążki interferencyjne. Jeśli na płytkę pada lekko rozbieżna wiązka promieni, to zmianę różnicy dróg optycznych można wyznaczyć licząc przesuwające się przez środek pola widzenia lunetki prążki interferencyjne (równego nachylenia).



Rys. 25.8

Ustawiając na drodze jednej z wiązek ciało przezroczyste o długości l i szukanym współczynniku załamania n (współczynnik załamania otaczającego ośrodka n_0) wywołujemy powstanie dodatkowej różnicy dróg optycznych równej $l(n-n_0)$, wskutek czego przez środek pola widzenia lunetki przechodzi $l(n-n_0)/\lambda$ prążków. Metodą tą można osiągać dokładność pomiaru różnicy współczynników załamania do $0{,}5 \cdot 10^{-6}$.

Na zakończenie wspomnimy jeszcze krótko o zastosowaniu zjawiska interferencji w przemyśle optycznym. Wykorzystuje się tam prążki interferencyjne równej grubości do badania dokładności oszlifowania powierzchni płytek szklanych. W tym celu na płytkę wzorcową o idealnie gładkiej powierzchni nakłada się płytkę badaną i ściska się je razem przy jednej z krawędzi. Wywołuje się w ten sposób powstanie „powietrznego klina" między płytkami. Obserwacja płytek w świetle odbitym wykazuje pojawienie się prążków interferencyjnych równej grubości. Jeśli mają one kształt pasków równoległych do ściskanej krawędzi, to płytka badana jest dobrze oszlifowana. Jeśli jednak kształt prążków odbiega od prostoliniowego, jest to dowód występowania nierówności na powierzchni płytki. Na rysunku 25.8 widzimy układ prążków w przypadku płytki dobrze oszlifowanej i mającej defekty powierzchniowe.

Z innych przyrządów, których działanie opiera się na zjawisku interferencji, wymienimy przykładowo: mikrointerferometry, mikroskopy interferencyjne, interferometry wielowiązkowe itp.

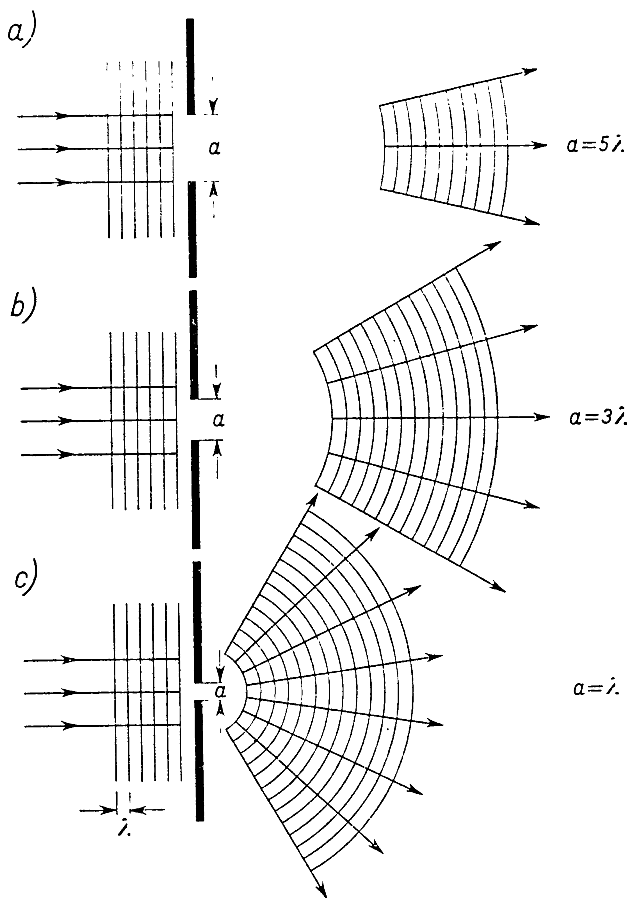
25.6. Ugięcie światła

W paragrafie 24.1 mówiliśmy o tym, że niemożliwe jest otrzymanie pojedynczego promienia świetlnego, czyli nieskończenie wąskiej wiązki promieniowania. Omówimy teraz to zagadnienie nieco dokładniej. Na rysunku 25.9 widzimy przesłonę z otworem,

którego rozmiary można stopniowo zmniejszać. Na kolejnych rysunkach a, b, c szerokość szczeliny a maleje w ten sposób, że

$$a_1 = 5\lambda, \quad a_2 = 3\lambda, \quad a_3 = 1\lambda.$$

Na szczelinę pada fala płaska. Poza szczelinę przedostaje się fala o powierzchni lekko zakrzywionej, rozchodząca się w obszarze szerszym niż wynikałoby to z prawa prostoliniowego rozchodzenia się światła. To „poszerzenie" wiązki świetlnej (lub dowolnej fali innego typu) jest właśnie istotą zjawiska ugięcia, czyli dyfrakcji. Przykłady ugięcia fal głosowych i fal na wodzie omawialiśmy już poprzednio (§ 11.4).

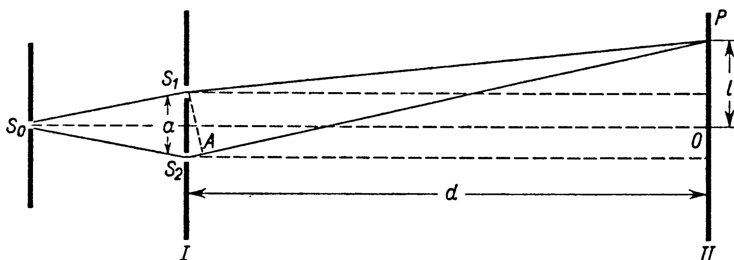


Rys. 25.9

Z rysunków 25.9a, b, c można wysnuć dalsze wnioski. Efekt ugięcia jest tym wyraźniejszy, im mniejsza jest szerokość szczeliny w stosunku do długości fali. W dziedzinie np. fal głosowych mamy do czynienia z długościami fal około 0,02–20 m, a zatem takie często spotykane „otwory", jak otwarte okno, drzwi itp., dają wyraźne efekty ugięciowe. W dziedzinie fal świetlnych mamy inne rzędy wielkości długości fali. Dłu-

gości granicznych fal widzialnych: czerwonej i fioletowej wynoszą około $4 \cdot 10^{-7}$ m i $8 \cdot 10^{-7}$ m. Nic więc dziwnego, że efekty zjawiska ugięcia rzadziej obserwujemy w dziedzinie światła widzialnego niż w dziedzinie fal akustycznych. Przy spełnieniu warunku, że rozmiary szczeliny są znacznie większe od długości fali, czyli $a \gg \lambda$, wpływ ugięcia jest do zaniedbania i można przyjmować, że światło rozchodzi się po liniach prostych, tzn. traktować zjawiska z punktu widzenia optyki geometrycznej. Gdy zaś warunek $a \gg \lambda$ nie jest spełniony, trzeba już uwzględniać naturę falową światła. Wtedy zjawiska są opisywane z punktu widzenia optyki falowej.

Interesującym faktem jest to, że zjawisko ugięcia światła znane było zarówno Huygensowi (1629–1695), jak i Newtonowi (1642–1727), ale żaden z nich nie umiał z niego skorzystać jako z doświadczalnego kryterium przemawiającego za lub przeciw jego teorii. Dopiero tacy badacze jak Fresnel, Fraunhofer, Young zjawisko to zbadali i wyjaśnili, o czym będziemy mówili w następnych punktach tego rozdziału.



Rys. 25.10

25.6.1. *Doświadczenie Younga.* Jeśli źródło światła umieścimy przed wąską szczeliną, to dzięki zjawisku ugięcia szczelinę można traktować jako źródło. Postępujemy tak w przypadku spektrometru, gdzie oświetlona szczelina kolimatora jest uważana za źródło światła. Podobnie postępował Young przeprowadzając w r. 1801 doświadczenie, w którym badał ugięcie i interferencję światła. Schemat tego doświadczenia przedstawiony jest na rys. 25.10. W nieprzezroczystym ekranie *I* wycięte są dwie wąskie szczeliny S_1 i S_2 prostopadłe do płaszczyzny rysunku, odległe od siebie o a. W odległości d znajduje się drugi ekran *II*, na którym badamy wyniki ugięcia i interferencji. Szczeliny S_1 i S_2 są oświetlone światłem pochodzącym ze źródła S_0 (będącego też małą szczeliną uginającą światło) ustawionego symetrycznie względem obu szczelin. Dzięki zjawisku ugięcia promienie rozchodzące się poza szczelinami S_1 i S_2 biegną we wszystkich kierunkach: możemy zatem szczeliny te traktować jako spójne źródła światła.

Rozważmy początkowo wyniki interferencji otrzymane na ekranie *II* przy użyciu światła monochromatycznego. Do punktu *O*, położonego symetrycznie względem szczelin, dochodzą fale w fazach zgodnych, nakładają się i powodują wystąpienie wzmocnienia światła. Szczelinom prostopadłym do płaszczyzny rysunku odpowiada na ekranie prążek jasny, przechodzący przez punkt *O*, również prostopadły do płaszczyzny rysunku.

Do innych punktów ekranu II dochodzą fale od szczelin S_1 i S_2 po przebyciu różnych dróg. Różnice tych dróg optycznych decydują o różnicy faz interferujących fal. Różnica dróg do punktu P równa się w przybliżeniu odcinkowi $S_2A = \delta$, powstałemu po wykreśleniu prostopadłej S_1A. Obliczamy wielkość δ wprowadzając oznaczenie $PO = l$:

$$S_2P^2 = d^2 + \left(l + \frac{a}{2}\right)^2,$$

$$S_1P^2 = d^2 + \left(l - \frac{a}{2}\right)^2.$$

Odejmujemy te równania stronami:

$$S_2P^2 - S_1P^2 = 2la,$$

skąd

$$(S_2P + S_1P)(S_2P - S_1P) = 2la.$$

Przy założeniu, że odległość d między ekranami jest bardzo duża w porównaniu z odległością a między szczelinami i odległością badanego punktu na ekranie od punktu centralnego O, można uważać, że

$$S_1P = S_2P = d,$$

wówczas

$$2d\delta = 2la,$$

$$\delta = \frac{la}{d}.$$

Zależność ta wynikła z rozważań czysto geometrycznych. Rozpatrzymy teraz zagadnienie z punktu widzenia fizyki i poszukamy odpowiedzi na pytanie, jaki warunek musi spełniać δ, aby punkt P leżał na: 1) prążku jasnym lub 2) prążku ciemnym. Dla prążków jasnych różnica dróg δ musi być równa całkowitej wielokrotności λ, czyli

$$\frac{l_m a}{d} = m\lambda, \quad \text{gdzie} \quad m = 1, 2, 3, \ldots,$$

tzn. kolejne prążki jasne tworzą się w odległości l_m od prążka centralnego:

$$l_m = m\frac{\lambda d}{a}, \quad \text{gdzie} \quad m = 1, 2, 3, \ldots \tag{25.4}$$

Prążki ciemne powstają wtedy, gdy odległość δ równa się nieparzystej wielokrotności $\lambda/2$, czyli

$$\frac{l_k a}{d} = (2k+1)\frac{\lambda}{2}, \quad \text{gdzie} \quad k = 0, 1, 2, 3, \ldots$$

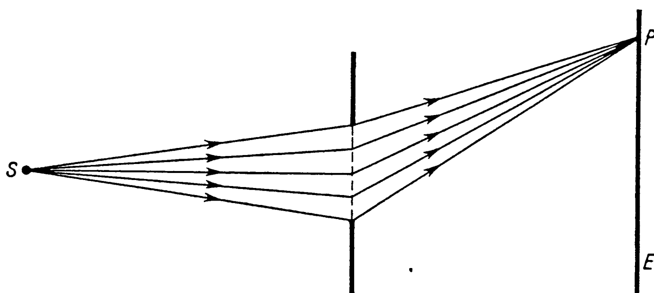
Ich odległość l_k od prążka centralnego wynosi

$$l_k = (2k+1)\frac{\lambda}{2}\frac{d}{a}, \quad \text{gdzie} \quad k = 0, 1, 2, 3, \ldots \tag{25.5}$$

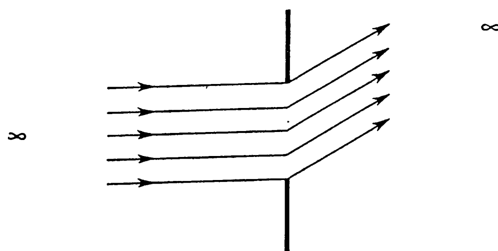
Jak widać z tych wzorów, odległość kolejnych prążków jasnych od prążka centralnego jest funkcją długości fali. Jeśli więc zamiast światła monochromatycznego użyje-

my źródła światła białego, to prążki jasne różnych barw dla danego m nie wypadną w tym samym miejscu na ekranie, a zatem nie wytworzą prążka białego. W tym przypadku powstają prążki barwne rozmieszczone na ekranie w ten sposób, że prążki odpowiadające krótszym falom znajdują się bliżej białego prążka centralnego. Otrzymujemy w ten sposób widmo. W zależności od wartości $m = 1, 2, 3, \ldots$ mówimy o widmie pierwszego, drugiego itd. rzędu. Dla większych wartości m obraz interferencyjny się zaciera, gdyż prążki interferencyjne krótkofalowe wyższych rzędów nakładają się na prążki interferencyjne długofalowe niższych rzędów.

Doświadczenie Younga umożliwia obliczanie długości fali. Mierzymy wielkości a, d, l i obliczamy długość fali λ dla prążka jasnego. Jednakże dokładność pomiaru λ tą metodą jest niewielka.



Rys. 25.11



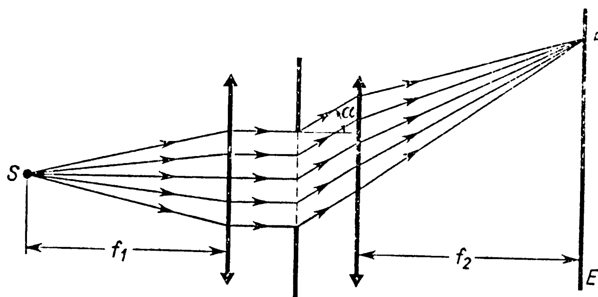
Rys. 25.12

25.6.2. *Ogólne wiadomości o dyfrakcji Fresnela i Fraunhofera.* W zależności od warunków, w jakich powstaje zjawisko ugięcia (czyli dyfrakcji), mówimy o dyfrakcji Fresnela lub o dyfrakcji Fraunhofera.

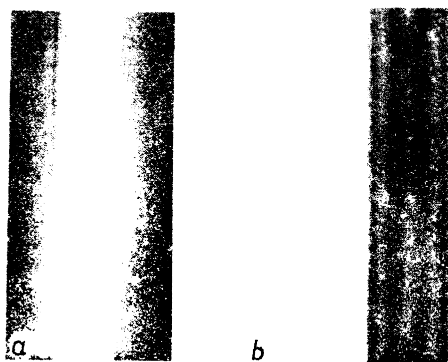
Z dyfrakcją Fresnela mamy do czynienia wtedy, gdy wiązka rozbieżna, pochodząca z bliskiego źródła, przechodzi przez otwór uginający światło lub obok zasłony uginającej. Wynik ugięcia obserwowany jest na ekranie umieszczonym w skończonej odległości od ciała uginającego światło. Schemat biegu promieni w takim przypadku przedstawiony jest na rys. 25.11.

Dyfrakcja Fraunhofera występuje wtedy, gdy na ciało uginające pada wiązka promieni równoległych, a więc wiązka pochodząca ze źródła nieskończenie odległego. Wynik ugięcia badany jest również na nieskończenie odległym ekranie (rys. 25.12).

W praktyce dyfrakcja Fraunhofera sprowadza się do badania ugięcia wiązek równoległych, uzyskanych za pomocą soczewek zbierających, gdy źródło światła umieszczone jest w ognisku soczewki. Podobnie równoległe wiązki ugięte skupiamy w płaszczyźnie ogniskowej soczewki zbierającej. Dzięki temu miejsce obserwacji przeniesione jest z nieskończoności do płaszczyzny ogniskowej drugiej soczewki (rys. 25.13). Teoretyczne wyjaśnienie rozmaitych efektów zjawiska ugięcia wiąże się z zasadą Huygensa (§ 11.4), opracowaną następnie matematycznie przez Fresnela i w związku z tym często nazywaną zasadą Huygensa–Fresnela. Nie wnikając w istotę rozważań teoretycznych omówimy kilka wyników doświadczalnych.



Rys. 25.13



Rys. 25.14

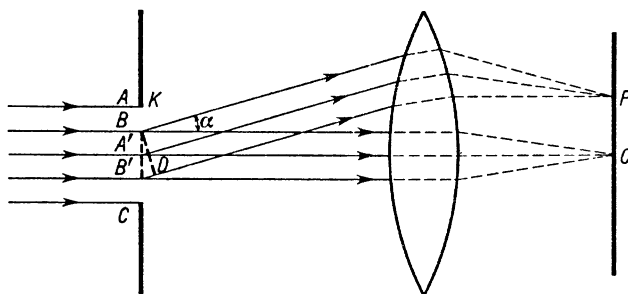
Dyfrakcja Fresnela. a) Ugięcie światła monochromatycznego na pojedynczej szczelinie: na ekranie za szczeliną otrzymujemy układ prążków ciemnych i jasnych. Prążek centralny jest jasny lub ciemny, zależnie od odległości ekranu od szczeliny (rys. 25.14b). b) Ugięcie na wąskiej przesłonie (np. na igle). Na ekranie powstaje układ prążków ciemnych i jasnych, ale tym razem prążek centralny jest zawsze jasny, niezależnie od odległości ekranu od zasłony (rys. 25.14a). c) Ugięcie na dwóch szczelinach (doświadczenie Younga) omówiono już szczegółowo w punkcie 25.6.1.

Dyfrakcja Fraunhofera. Wiąże się ona z działaniem jednego z podstawowych urządzeń optyki falowej, a mianowicie z siatką dyfrakcyjną. W punktach 25.6.3 i 25.6.4 omówimy bardziej szczegółowo dwa przypadki, a mianowicie:

1) ugięcie na pojedynczej szczelinie o szerokości a, porównywalnej z długością fali,

2) ugięcie na siatce dyfrakcyjnej.

25.6.3. *Dyfrakcja Fraunhofera na pojedynczej szczelinie.* Na szczelinę K (rys. 25.15) pada fala płaska, przedstawiona w postaci wiązki promieni równoległych. Poza szczeliną K światło rozchodzi się nie tylko w kierunku przedłużenia promieni pierwotnych, lecz także dzięki ugięciu odchyla się od pierwotnego kierunku o kąty zawarte w granicach od 0° do 90°. Każdy punkt szczeliny ugina światło. Na rysunku poza szczeliną zaznaczony jest bieg promieni nieugiętych oraz bieg trzech promieni ugiętych pod kątem α względem kierunku pierwotnego. W pewnej odległości za ekranem ze szczeliną ustawiona jest soczewka zbierająca, która skupia wiązki promieni równoległych w płaszczyźnie ogniskowej. Promienie nieugięte tworzą obraz w punkcie O. Wiązki równoległe, ugięte pod różnymi kątami, po załamaniu w soczewce przecinają się w różnych punktach płaszczyzny ogniskowej soczewki (powyżej i poniżej punktu O). W wyniku interferencji powstają tam wzmocnienia, osłabienia lub zupełne wygaszenia światła. Ustal-



Rys. 25.15

my warunek, jaki musi spełniać kąt ugięcia, aby powstało pierwsze (licząc od punktu O) wygaszenie światła. W tym celu podzielmy w myśli szczelinę na dwie równe części. W obu połówkach szczeliny wyodrębniamy parę tzw. odpowiadających sobie punktów: najwyższemu punktowi A górnej połowy odpowiada najwyższy punkt A' dolnej części. Środkowemu punktowi B górnej połowy odpowiada punkt B' leżący w środku połówki dolnej itd. Z promieni równoległych (rys. 25.15), padających na szczelinę, wyodrębniamy dwa promienie przechodzące przez parę odpowiadających sobie punktów, np. B i B'. Poza szczeliną rysujemy promienie nieugięte, biegnące od B i B' oraz promienie ugięte pod kątem α. Promienie ugięte (równoległe) załamują się w soczewce i wytwarzają obraz w punkcie P. Można udowodnić, że drogi optyczne* wszystkich promieni ugiętych pod kątem α, liczone od dowolnej płaszczyzny prostopadłej do kierunku ich biegu do punktu ich zbieżności P za soczewką, są jednakowe. A zatem promień ugięty

* Przypominamy, że droga optyczna na ogół nie równa się drodze geometrycznej promienia. Miarą drogi optycznej jest iloczyn współczynnika załamania danego ośrodka i długości drogi geometrycznej, czyli

$$l_{opt} = n l_{geom}$$

w punkcie B' przebywa do punktu P drogę optyczną dłuższą od drogi promienia ugiętego w punkcie B o odcinek $B'D = \delta$. Jeśli $\delta = \lambda/2$, to w wyniku interferencji w punkcie P będzie ciemno. Będzie to pierwszy ciemny obraz szczeliny w otoczeniu jasnego, centralnego obrazu w punkcie O.

Z zależności geometrycznych wynika, że

$$\delta = \frac{a}{2}\sin\alpha,$$

gdzie a równa się szerokości szczeliny. Stąd warunek dla kąta ugięcia α_1, odpowiadającego pierwszemu prążkowi ciemnemu

$$\sin\alpha_1 = \frac{\lambda}{a}.$$

Chwila zastanowienia wystarczy do stwierdzenia, że taki sam warunek wynika z rozważania ugięcia na dowolnie wybranej innej parze odpowiadających sobie punktów szczeliny. Jest to więc ogólny warunek charakteryzujący powstawanie pierwszego prążka ciemnego. Chcąc ustalić warunek powstawania dalszych prążków ciemnych należałoby podzielić szczelinę na większą, parzystą liczbę równych części, np. na 4, 6 itd. Biorąc pary odpowiadających sobie punktów, np. na dwóch sąsiednich ćwiartkach szczeliny, otrzymalibyśmy warunek wygaszania światła przy interferencji w postaci

$$\delta = \frac{a}{4}\sin\alpha_2, \qquad \sin\alpha_2 = \frac{\lambda}{2}\frac{4}{a} = \frac{2\lambda}{a}.$$

Kąt α_2 oznacza kąt ugięcia odpowiadający wytwarzaniu się drugiego prążka ciemnego. Działanie obu par kolejnych ćwiartek wzajemnie się znosi. Dla dalszych prążków ciemnych obowiązuje warunek

$$\sin\alpha_m = m\frac{\lambda}{a}, \qquad \text{gdzie } m = 1, 2, 3, \ldots$$

Warunek powstawania prążka jasnego znajdziemy dzieląc szczelinę na nieparzystą liczbę równych części, np. na 3. Jeśli dwie kolejne części wygaszą wzajemnie swoje działanie, to pozostanie niewygaszone działanie części trzeciej. A zatem warunek powstania pierwszego, poza centralnym, prążka jasnego przyjmie postać

$$\frac{a}{3}\sin\alpha_1 = \frac{\lambda}{2}, \qquad \text{czyli } \sin\alpha_1 = \frac{3\lambda}{2a}.$$

Ogólnie biorąc, k-ty prążek jasny powstanie przy ugięciu pod kątem α_k spełniającym równanie

$$\sin\alpha_k = \frac{(2k+1)}{a}\frac{\lambda}{2}, \qquad \text{gdzie } k = 1, 2, 3, \ldots \tag{25.6}$$

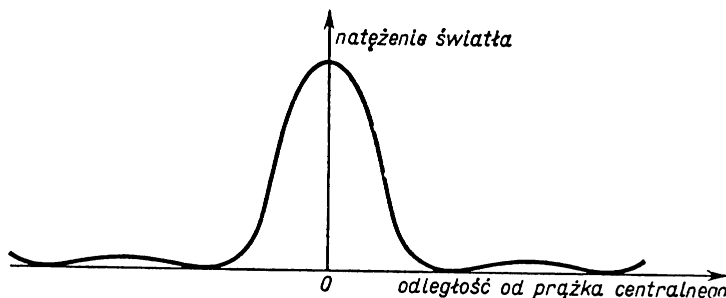
Obraz uzyskany na ekranie jest symetryczny względem jasnego prążka centralnego. Ten środkowy prążek jest najjaśniejszy i dość szeroki, dalsze prążki jasne są znacznie słabsze, lecz również szerokie, rozmyte. Przedstawia to wykres na rys. 25.16.

Ze wzoru (25.6) określającego kąty ugięcia odpowiadające prążkom jasnym wynika:

1) zależność kąta ugięcia od długości fali: ze wzrostem długości fali rośnie kąt ugięcia, a ze wzrostem kąta ugięcia rośnie odległość prążka badanego od prążka centralnego.

2) zależność kąta ugięcia od szerokości szczeliny.

Z zależności 1) wynika możliwość otrzymania na ekranie widma. W przypadku stosowania kolejno różnych źródeł światła monochromatycznego można stwierdzić, że prążki fioletowe pierwszego rzędu tworzą się bliżej prążka centralnego niż prążki tegoż rzędu o barwach pozostałych. Przy użyciu światła białego otrzymuje się widmo ciągłe.



Rys. 25.16

Z zależności drugiej wnioskujemy, że rozsunięcie prążków w stosunku do prążka centralnego staje się większe, jeśli nie zmieniając źródła zmienimy szczelinę na węższą.

25.6.4. *Siatka dyfrakcyjna.* Od zjawisk ugięcia i interferencji, związanych z pojedynczą szczeliną, przechodzimy do zjawisk tegoż typu występujących przy przejściu światła przez tzw. siatki dyfrakcyjne.

Siatka dyfrakcyjna jest to zbiór równoległych do siebie szczelin przepuszczających światło, rozmieszczonych w jednakowych odstępach. Siatkę dyfrakcyjną można uzyskać kreśląc ostrzem diamentu szereg równoległych rys, np. na powierzchni szkła w równych odstępach od siebie. Rysy stanowią te części powierzchni szkła, które nie przepuszczają światła. „Szczelinami" przepuszczającymi światło są powierzchnie szkła między rysami. Dobre siatki mają kilkaset (do 2000) rys na milimetrze. Na siatkę dyfrakcyjną rzucamy wiązkę promieni równoległych (dyfrakcja Fraunhofera). Obraz dyfrakcyjno-interferencyjny otrzymany na ekranie za pomocą soczewki zbierającej jest wynikiem nałożenia dwóch efektów, a mianowicie:

1) każda szczelina z osobna daje taki obraz dyfrakcyjny, o jakim mówiliśmy poprzednio w punkcie 25.6.3,

2) wszystkie szczeliny działając łącznie dają wspólny obraz interferencyjny.

Nie powtarzając już tego, co mówiliśmy o działaniu pojedynczej szczeliny, przejdziemy od razu do rozpatrzenia współdziałania poszczególnych szczelin. Niech szerokość każdej szczeliny wynosi b, a szerokość odstępu między szczelinami — c (rys. 25.17). Wyodrębniając, podobnie jak poprzednio, pary odpowiadających sobie punktów na sąsiednich szczelinach, znajdziemy, że ich wzajemne odległości wynoszą

$$d = b+c.$$

Wielkość d nazywamy stałą siatki dyfrakcyjnej.

Zbadajmy wiązkę ugiętą pod kątem α na odpowiadających sobie punktach sąsiednich szczelin. Jeśli różnica dróg sąsiednich promieni do miejsca spotkania na ekranie za soczewką (nie zaznaczone na rys. 25.17)

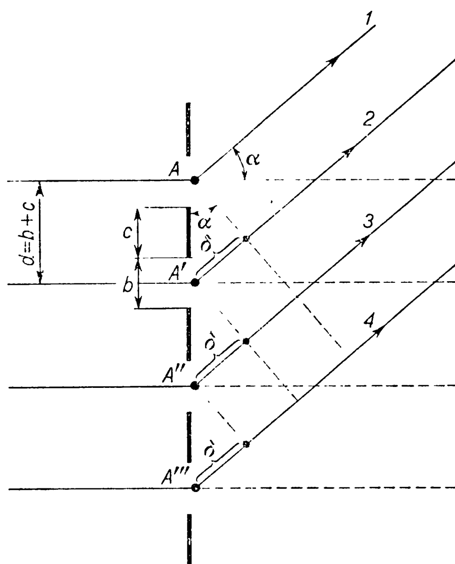
$$\delta = \lambda m$$

(gdzie m oznacza liczbę całkowitą), to promienie *1* i *2* wzmacniają się przy interferencji. Podobnie wzmacniają się promienie *2* i *3*, *1* i *3* itd. Ale z zależności geometrycznych wynika, że

$$\delta = d \sin \alpha .$$

A zatem warunek wytworzenia prążka jasnego przy współdziałaniu wszystkich szczelin można wyrazić w postaci wzoru

$$d \sin \alpha_m = m\lambda \quad \text{dla} \quad m = 0, 1, 2, 3, \ldots \tag{25.7}$$



Rys. 25.17

Wartość $m = 0$ odpowiada centralnemu prążkowi jasnemu na ekranie. Zastanówmy się nad ilością energii promienistej, dochodzącej do tego centralnego prążka jasnego. Od każdej ze szczelin dochodzi fala o amplitudzie A. Wszystkie fale spotykają się w fazach zgodnych. Amplituda fali wypadkowej jest sumą amplitud składowych, czyli równa się nA, gdzie n oznacza liczbę szczelin. Energia fali jest proporcjonalna do kwadratu amplitudy, a zatem do n^2A^2. Natężenie promieniowania w prążku centralnym jest n^2 razy większe niż w przypadku pojedynczej szczeliny.

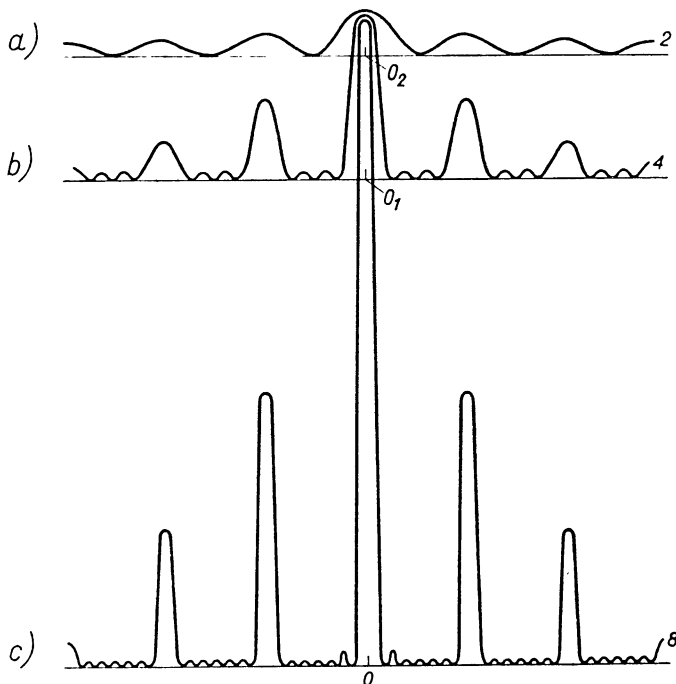
Liczba szczelin siatki dyfrakcyjnej wpływa nie tylko na wystąpienie mocniejszych i nowych maksimów światła. Wskutek interferencji powstają również dodatkowe minima światła. Wynik użycia siatki dyfrakcyjnej jest taki, że na ekranie zamiast rozmytych jasnych prążków powstają prążki ostre i bardzo jasne, poprzedzielane prążkami

ciemnymi. Na rys. 25.18a, b, c widzimy zmianę charakteru otrzymywanego obrazu. Rzędne punktów nakreślonych krzywych przedstawiają natężenia promieniowania, a odcięte (liczone odpowiednio od punktów O, O_1 lub O_2) — odległości od prążka centralnego. Kolejne rysunki odnoszą się do przypadków zastosowania dwóch, czterech i ośmiu szczelin.

Równanie

$$d \sin \alpha_m = m\lambda,$$

określające powstawanie widma m-tego rzędu przy użyciu siatki dyfrakcyjnej o stałej d, można stosować do obliczania długości fali. Do tego celu używany jest zwykle spektrometr lub spektrograf siatkowy. Budowę tych przyrządów omówimy dokładnie w rozdz. 28.



Rys. 25.18

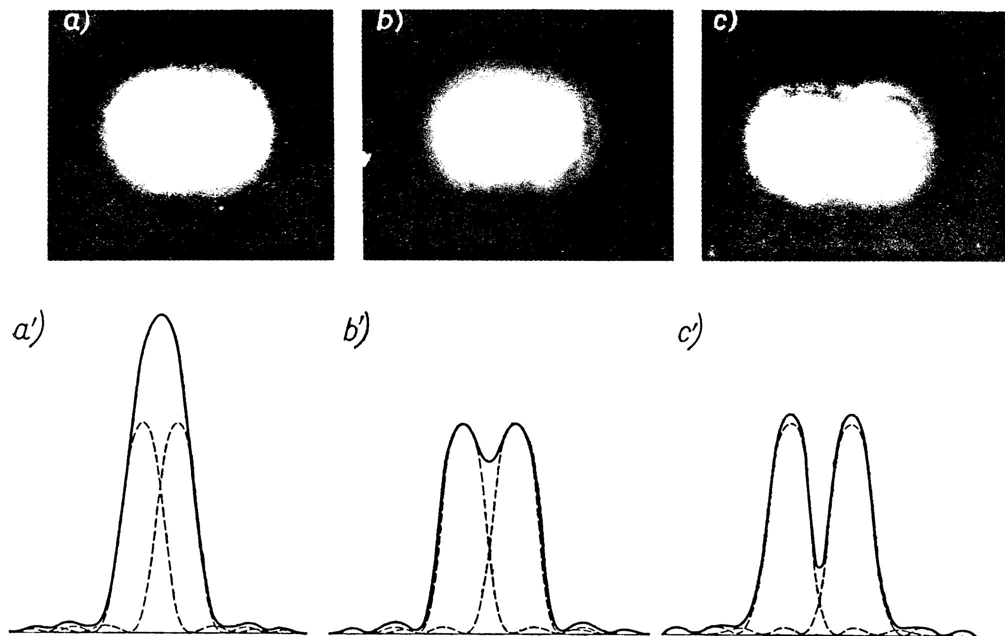
25.7. Zdolność rozdzielcza przyrządów soczewkowych

Wiemy już, że umieszczenie punktowego źródła światła przed szczeliną powoduje powstanie obrazu dyfrakcyjnego w postaci na przemian jasnych i ciemnych prążków. Ustawienie źródła przed małym kołowym otworem wywołuje powstanie obrazu dyfrakcyjnego w postaci rozmytego krążka otoczonego na przemian jasnymi i ciemnymi pierścieniami.

Każdy soczewkowy przyrząd optyczny ma jakiś otwór wejściowy, którego rozmiary wyznacza bezpośrednio średnica obiektywu lub też średnica otworu zastosowanej kołowej przesłony wejściowej. A zatem obraz każdego punktu świecącego przedmiotu

rozmywa się w krążek otoczony koncentrycznymi pierścieniami. Można wykazać, że rozmiary kątowe ($\theta \approx \sin\theta$) pierwszego minimum dyfrakcyjnego powstającego wskutek dyfrakcji Fraunhofera na otworze kołowym o średnicy D promieni rozchodzących się z punktowego źródła wynoszą $1{,}22\frac{\lambda}{D}$ (czyli są proporcjonalne do długości fali użytego światła i odwrotnie proporcjonalne do średnicy otworu D). Promień pierwszego pierścienia ciemnego powstającego w płaszczyźnie ogniskowej równa się $1{,}22\frac{\lambda f}{D}$.

Jeżeli weźmiemy pod uwagę dwa bliskie siebie punkty świecące, to każdemu z nich będzie odpowiadał obraz dyfrakcyjny w płaszczyźnie ogniskowej soczewki. Rysunek 25.19a dotyczy tak małej odległości tych dwóch punktów, że są one nierozróżnialne: obrazy dyfrakcyjne nakładają się prawie dokładnie. Przebieg krzywej rozkładu natężeń w wypadkowym obrazie ma charakter ogólny zbliżony, jak to widać z rysunku a′, do przebiegu krzywych rozkładu natężeń w poszczególnych obrazach.



Rys. 25.19

Rysunki 25.19b i b′ dotyczą takiej odległości badanych dwóch punktów, że ich obrazy dyfrakcyjne są ledwo rozróżnialne, a krzywe rozkładu natężeń są tak względem siebie przesunięte, że maksimum jednej z nich przypada na pierwsze minimum drugiej. O takich obrazach mówimy, że spełniają kryterium Rayleigha. Rysunki 25.19c i c′ dotyczą jeszcze większej odległości między badanymi punktami, przy której obrazy są już wyraźnie rozróżnialne.

Inaczej mówiąc, jeśli środki obrazów dyfrakcyjnych odpowiadających dwóm bliskim

punktom źródła światła leżą w odległości mniejszej od promienia jasnego krążka dyfrakcyjnego, to wspomnianych punktów nie jesteśmy w stanie rozdzielić: zlewają się one w obrazie w jedną całość.

Zagadnienie zdolności rozdzielczej mikroskopu rozważymy w nieco inny, przybliżony sposób z tego względu, że warunki towarzyszące tworzeniu się obrazu w mikroskopie odbiegają od warunków dyfrakcji Fraunhofera.

25.7.1. *Zdolność rozdzielcza mikroskopu.* Rozważmy jeszcze raz tworzenie się obrazów w mikroskopie. Przedmiot oglądany umieszczony jest na szkiełku przedmiotowym i silnie oświetlony od dołu za pomocą wklęsłego zwierciadełka, skierowującego światło na przedmiot. Promienie świetlne, przechodząc przez preparat przezroczysty, ulegają ugięciu i trafiają do obiektywu mikroskopu. Dla uproszczenia zagadnienia potraktujemy preparat oglądany jak siatkę dyfrakcyjną. Jeżeli rozwartość obiektywu jest taka, że trafiają do niego tylko promienie osiowe nieugięte, to otrzymujemy w polu widzenia okularu tylko mniej więcej jednolicie oświetlone jasne tło. Nie mamy możliwości rozróżniania żadnych szczegółów budowy przedmiotu. Tylko wtedy dostrzeżemy szczegóły budowy preparatu, gdy do obiektywu mikroskopu trafia prócz wiązki nieugiętej przynajmniej jedna wiązka ugięta, wytwarzająca obraz jasny pierwszego rzędu.

Na rysunku 25.20a przedstawiony jest bieg wiązki osiowej i pierwszej wiązki ugiętej pod kątem α, spełniającym warunek siatki dyfrakcyjnej

$$d \sin\alpha = \lambda m,$$

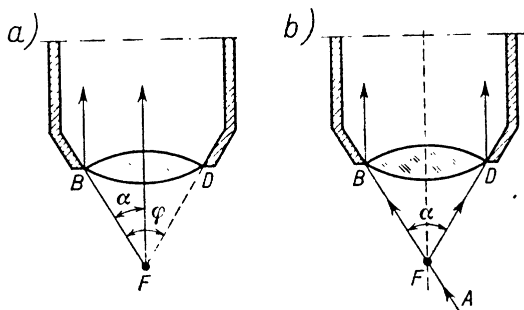
gdzie $m = 1$. Kąt α występujący w tym wzorze jest miarą połowy rozwartości kątowej φ obiektywu, zwanej też kątem aperturowym lub krócej — aperturą obiektywu. *Rozwartość kątowa* (*kąt aperturowy*) obiektywu zdefiniowana jest jako kąt zawarty między skrajnymi promieniami wchodzącymi do obiektywu, pochodzącymi z punktu przedmiotu obranego na osi optycznej. Z pewnym przybliżeniem można przyjmować, że wybrany punkt przedmiotu leży w ognisku obiektywu.

W podanym wyżej warunku wielkość d odpowiada odległości dwóch punktów preparatu, które w obserwacji mikroskopowej odróżniamy od siebie, czyli widzimy oddzielnie dzięki wystąpieniu obrazu dyfrakcyjnego pierwszego rzędu. Punktów leżących bliżej siebie nie odróżniamy, gdyż odpowiada im większy kąt ugięcia związany z widmem pierwszego rzędu. Promienie ugięte pod takim kątem już nie trafiają do obiektywu. Może warto tu przypomnieć, że dla przeciętnego oka zdolność dostrzegania szczegółów bez pomocy dodatkowych urządzeń optycznych ograniczona jest kątem widzenia równym jednej minucie łuku, co w skali liniowej odpowiada mniej więcej $\frac{1}{15}$ mm przy obserwacji z odległości wyraźnego widzenia.

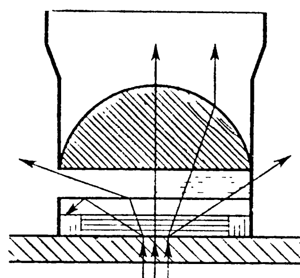
Określanie możliwości dostrzegania szczegółów w przypadku użycia mikroskopu przez podawanie minimalnej odległości punktów, które jeszcze rozróżniamy przy obserwacji mikroskopowej, jest niewygodne. Zastąpiono je podawaniem wielkości zwanej zdolnością rozdzielczą mikroskopu. Przez *zdolność rozdzielczą* mikroskopu rozumiemy odwrotność najmniejszej odległości dwóch punktów, które jeszcze rozróżniamy, czyli

$$\frac{1}{d} = \frac{\sin\alpha}{\lambda}. \tag{25.8}$$

Zwiększenie zdolności rozdzielczej mikroskopu jest zagadnieniem bardzo ważnym. Im większa jest zdolność rozdzielcza, tym więcej szczegółów rozróżniamy na preparacie. Jak widać ze wzoru, mamy dwie drogi do zwiększenia zdolności rozdzielczej: 1) zwiększenie kąta rozwartości obiektywu lub 2) zmniejszenie długości fali użytego światła. Zwiększenie kąta α można uzyskać nie tylko przez wymianę danego obiektywu na obiektyw o większej rozwartości. Można też zastosować oświetlenie boczne preparatu zamiast oświetlenia osiowego, tak jak to przedstawiono na rys. 25.20b. Kierunek AB odpowiada biegowi wiązki centralnej padającej na preparat, kierunek FD przedstawia bieg granicznej wiązki ugiętej, wpadającej jeszcze do obiektywu. Jak widać z rysunku, otrzymujemy w ten sposób dwukrotne zwiększenie kąta α w porównaniu z zastosowaniem oświetlenia osiowego.



Rys. 25.20


Rys. 25.21

Stosując światło fioletowe otrzymamy największą zdolność rozdzielczą przy obserwacji wizualnej za pomocą zwykłego mikroskopu. Można też stosować promieniowanie nadfioletowe, obserwację wizualną trzeba jednak wtedy zastąpić obserwacją fotograficzną.

Mikroskopy tzw. immersyjne mają większe zalety niż mikroskopy zwykłe. Pozwalają one na uzyskanie obrazów o większej jasności i mają większą zdolność rozdzielczą niż mikroskopy zwykłe. W mikroskopach immersyjnych przestrzeń między szkiełkiem nakrywkowym i obiektywem wypełniona jest tzw. cieczą immersyjną, tzn. cieczą o możliwie dużym współczynniku załamania. Jako ciecze immersyjne stosuje się często olejek cedrowy (współczynnik załamania $n = 1{,}5$) oraz monobromonaftalen (współczynnik załamania $n = 1{,}66$).

Aby wyjaśnić przyczyny powstawania jaśniejszego obrazu w mikroskopie immersyjnym, prześledzimy na tym samym rysunku bieg promieni przy użyciu cieczy immersyjnej (prawa strona rys. 25.21) i bez niej (lewa strona rysunku). W celu uproszczenia zagadnienia załóżmy, że współczynnik załamania cieczy immersyjnej równa się współczynnikowi załamania szkła.

Przy przejściu promienia ze szkła nakrywkowego do cieczy immersyjnej i następnie z cieczy do obiektywu promień nie ulega załamaniu, co wynika z wprowadzonego upraszczającego założenia. Natomiast przy przejściu ze szkła nakrywkowego do po-

wietrza promień ulega załamaniu i odchyla się od osi przyrządu. Tym samym zmniejsza się liczba promieni wpadających do obiektywu, co decyduje o mniejszej jasności powstającego obrazu. Oczywiście, w przypadku gdy ciecz immersyjna ma współczynnik załamania większy od szkła, sytuacja przedstawia się jeszcze korzystniej, gdyż załamane w cieczy promienie odchylają się w stronę osi przyrządu i jeszcze większa ich liczba bierze udział w wytwarzaniu obrazu.

Zwiększenie zdolności rozdzielczej w mikroskopie immersyjnym wiąże się z tym, że długość fali nie jest stałą cechą promieniowania, lecz zmienia się w zależności od ośrodka, przez który przechodzi promieniowanie. Jeśli długość fali w próżni oznaczymy przez λ, a długość fali w ośrodku o współczynniku załamania n przez λ_1, to z określenia współczynnika załamania wynika, że

$$n = \frac{c}{v} = \frac{cT}{vT} = \frac{\lambda}{\lambda_1},$$

skąd

$$\lambda_1 = \frac{\lambda}{n}.$$

Innymi słowy, wprowadzenie cieczy immersyjnej o współczynniku załamania n ma skutek taki, jak gdybyśmy światło o długości fali λ zastąpili światłem o długości fali n razy mniejszej. Zdolność rozdzielcza mikroskopu immersyjnego wzrasta zatem n-krotnie w stosunku do zdolności rozdzielczej tego samego mikroskopu użytego bez cieczy immersyjnej.

Równanie określające zdolność rozdzielczą mikroskopu immersyjnego ma postać

$$\frac{1}{d} = \frac{n \sin\alpha}{\lambda}, \tag{25.9}$$

gdzie λ oznacza długość fali promieniowania stosowanego podczas obserwacji w odniesieniu do próżni. Wielkość $n \sin\alpha$ nazywamy *aperturą numeryczną* albo *aperturą liczbową obiektywu mikroskopu*. Aby przypomnieć znaczenie kąta α, należy powrócić do rys. 25.20a. Kąt α jest kątem, pod którym widać z preparatu (z ogniska) promień obiektywu. Zmiana obiektywu mikroskopu wywołuje zmianę apertury i zdolności rozdzielczej.

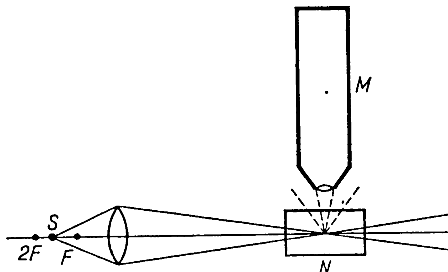
Stosując promieniowanie nadfioletowe i immersję można w nowoczesnych mikroskopach odróżniać punkty leżące w odległości 1200 A°. Większą zdolność rozdzielczą można uzyskać w mikroskopach elektronowych. Przy ich użyciu można rozróżniać punkty leżące w odległości 5–6 Å.

W technice mikroskopowej oprócz mikroskopów zwykłych i immersyjnych stosowane są jeszcze inne rodzaje, jak np. ultramikroskopy, mikroskopy interferencyjne i mikroskopy z kontrastem fazowym.

Ultramikroskopy mają inne zadania niż zwykłe mikroskopy optyczne. Służą one do stwierdzenia obecności bardzo drobnych cząstek, bez uwidaczniania szczegółów ich budowy. Schemat ustawienia ultramikroskopu przedstawiony jest na rys. 25.22.

Źródło światła S, umieszczone między punktami F i $2F$ soczewki zbierającej o osi optycznej poziomej, daje wiązkę zbieżną, skierowaną w stronę naczynia N, wypełnio-

nego przezroczystą cieczą. Zwykły mikroskop optyczny ustawiony jest pionowo, tak że jego oś przypada nad punktem zbieżności wiązki. Całość urządzenia znajduje się w zaciemnionym pokoju, a źródło światła jest tak osłonięte, że promienie bezpośrednio od niego idące nie wchodzą do obiektywu mikroskopu, zatem pole widzenia mikroskopu jest ciemne. Jeśli jednak na tym ciemnym tle spostrzeżemy jasne punkty, będzie to dowód, że w cieczy (w naczyniu) znajdują się cząsteczki zawiesin uginające światło.



Rys. 25.22

Część promieniowania ugiętego trafia do obiektywu mikroskopu, powodując pojawianie się jasnych punktów. Na podstawie takich obserwacji nie można określić kształtu cząstek uginających światło, można tylko stwierdzić ich obecność. Za pomocą ultramikroskopu można wykrywać cząstki o rozmiarach do 400 Å.

W rozdziale tym mówiliśmy przede wszystkim o mikroskopach optycznych. Mikroskop elektronowy omówimy w punkcie 30.7.1.

Pytania i zadania

1. Jakie można wymienić ważne momenty z historii rozwoju poglądów na naturę światła?
2. Podać bliższą charakterystykę światła jako fali elektromagnetycznej.
3. Z czym wiążą się trudności obserwowania zjawiska interferencji światła?
4. Na czym polega spójność promieni? Dlaczego promienie wychodzące z dwóch niezależnych źródeł nie mogą tworzyć wiązek spójnych?
5. W jaki sposób można otrzymać obserwowalną interferencję światła?
6. Jakie warunki muszą być spełnione, aby interferujące promienie: a) wzmacniały się, b) osłabiały, c) wygaszały?
7. Czy w przypadku ustalania różnicy faz interferujących promieni należy brać pod uwagę długość przebytych przez nie dróg geometrycznych, czy też dróg optycznych? Jak brzmi definicja drogi optycznej promienia?
8. Dzięki czemu zarówno w świetle przechodzącym, jak i w świetle odbitym od cienkich warstw obserwuje się efekty barwne? Podać przykłady.
9. Kiedy i w jaki sposób obserwuje się prążki interferencyjne: a) jednakowej grubości, b) jednakowego nachylenia?
10. Na czym polega doświadczenie z pierścieniami Newtona? Jakie może być ich zastosowanie? Jakie jeszcze inne metody można zastosować do tego samego celu?
11. Jaka jest budowa i działanie interferometru Michelsona? Jakie praktyczne zastosowanie ma ten przyrząd?
12. Wymienić kilka zastosowań zjawiska interferencji w metrologii.

13. Na czym polega zjawisko dyfrakcji (ugięcia)? W jakich warunkach należy się liczyć ze skutkami ugięcia?

14. Czym różnią się warunki wywoływania i obserwacji dyfrakcji Fresnela i Fraunhofera? Podać przykłady obu rodzajów dyfrakcji.

15. Dlaczego doświadczenie Younga nie stanowi dokładnej metody wyznaczania długości fali świetlnej?

16. Jakie dwa zjawiska przyczyniają się do powstawania ostatecznego obrazu w siatce dyfrakcyjnej? Na jakim wzorze opiera się wyznaczanie długości fali świetlnej za pomocą siatki? W skład jakiego przyrządu wchodzi siatka dyfrakcyjna?

17. Dlaczego przy tworzeniu się obrazów w przyrządach soczewkowych odgrywa zawsze rolę ugięcie światła?

18. Jak brzmi kryterium Rayleigha? Wyjaśnić jego sens fizyczny.

19. Co to jest zdolność rozdzielcza mikroskopu? Jakie czynniki wpływają na powiększenie zdolności rozdzielczej? Na czym polega różnica między mikroskopem immersyjnym i zwykłym mikroskopem optycznym? Co nazywamy aperturą numeryczną mikroskopu? Do czego służy ultramikroskop?

20. Jaką najmniejszą grubość może mieć płaska błonka mydlana, jeśli oglądana w świetle odbitym pod kątem $\alpha = 30°$ względem normalnej ma zabarwienie zielone ($\lambda_z = 5 \cdot 10^{-7}$ m). Współczynnik załamania błonki $n = 1{,}33$.

Odp. $1{,}04 \cdot 10^{-7}$ m.

21. Klin powietrzny o długości l i wysokości h wytworzony jest między dwiema idealnie oszlifowanymi płytkami szklanymi. Jaka jest odległość dwóch sąsiednich prążków interferencyjnych, jeśli kąt padania na klin promieni monochromatycznych o długości λ równa się α?

Odp. $\dfrac{l\lambda}{2h\cos\alpha}$.

22. Dwie jednakowe, gładko oszlifowane płytki szklane stykają się ze sobą wzdłuż boku o długości $l = 10$ cm. Przeciwległe boki są rozsunięte na odległość $d = 0{,}05$ mm (np. wskutek wsunięcia między nie cienkiej bibułki). Promienie o długości fali $\lambda = 6 \cdot 10^{-7}$ m padają pod kątem 60° na powierzchnię klina. Ile prążków interferencyjnych można zaobserwować na jednostce długości płytki?

Odp. 8,3 na 1 cm.

23. W doświadczeniu z pierścieniami Newtona zastosowano światło monochromatyczne sodowe o długości fali $\lambda = 5{,}89 \cdot 10^{-7}$ m i stwierdzono, że średnice dwóch sąsiednich jasnych pierścieni wynoszą 0,72 mm i 0,84 mm. Ile wynosi zdolność zbierająca zastosowanej szklanej soczewki ($n = 1{,}5$)?

Odp. 1,56 dioptrii.

24. Układ optyczny do badania pierścieni Newtona zanurzony jest w wodzie. Jaka jest odległość między trzecim i czwartym pierścieniem Newtona w świetle odbitym, jeśli promień krzywizny soczewki $R = 1$ m, współczynnik załamania wody $n = 1{,}33$, długość fali zastosowanego promieniowania $\lambda = 6 \cdot 10^{-7}$ m?

Odp. 0,19 mm.

25. Na szczelinę o szerokości $a = 0{,}5$ mm pada prostopadle wiązka promieni monochromatycznych. Na ekranie odległym o $d = 4{,}2$ m tworzy się obraz dyfrakcyjny, przy czym odległość pierwszego prążka ciemnego od środka obrazu dyfrakcyjnego wynosi $l = 4{,}2$ mm. Obliczyć długość fali zastosowanego promieniowania.

Odp. $5 \cdot 10^{-4}$ mm.

26. Wąska szczelina oświetlona jest równoległą wiązką światła białego padającego prostopadle. Dla jakiej długości fali środek trzeciego ciemnego prążka pokryje się ze środkiem drugiego ciemnego prążka, odpowiadającego $\lambda = 6{,}9 \cdot 10^{-7}$ m.

Odp. $4{,}6 \cdot 10^{-7}$ m.

27. Na siatkę dyfrakcyjną o 310 rysach na 1 milimetrze pada prostopadle równoległa wiązka światła białego. Obliczyć kąt ugięcia odpowiadający pokrywaniu się na ekranie widma rzędu m dla $\lambda_1 = 5{,}4 \cdot 10^{-7}$ m i widma rzędu $(m+1)$ dla $\lambda_2 = 4{,}05 \cdot 10^{-7}$ m.

Odp. 30°.

ROZDZIAŁ 26

Optyka falowa (II)
Polaryzacja światła

26.1. Istota zjawiska polaryzacji światła

Omawiane dotychczas zjawiska ugięcia i interferencji światła, aczkolwiek świadczą o jego naturze falowej, to jednak nie pozwalają wyciągać wniosków co do charakteru fali. Zjawiska te są bowiem charakterystyczne zarówno dla fal podłużnych, jak i dla fal poprzecznych. O poprzecznym charakterze fal świetlnych świadczy zjawisko polaryzacji.

Rozważając zjawisko polaryzacji światła należy pamiętać o tym, że fale świetlne — to fale elektromagnetyczne, związane z rozchodzeniem się w przestrzeni okresowo zmiennych pól: elektrycznego o natężeniu **E** i magnetycznego o natężeniu **H** (lub indukcji magnetycznej **B**).

Bardzo pożyteczne jest wprowadzenie do rozważań trzeciego, dodatkowego wektora **k**, o długości dowolnej, np. jednostkowej. Kierunek tego wektora jest zgodny z kierunkiem rozchodzenia się promienia świetlnego. Wektor ten zwany jest czasem *wektorem kierunkowym*. Badania kierunków wektorów **E**, **H** i **k** wykazały, że w przestrzeni nieograniczonej, jednorodnej $\mathbf{k} \perp \mathbf{E}$, $\mathbf{k} \perp \mathbf{H}$ i $\mathbf{E} \perp \mathbf{H}$. Właśnie w tej wzajemnej prostopadłości wektorów **E**, **H** i **k** tkwi poprzeczność fali świetlnej. Doświadczalnym dowodem poprzecznego charakteru fali świetlnej jest zjawisko polaryzacji.

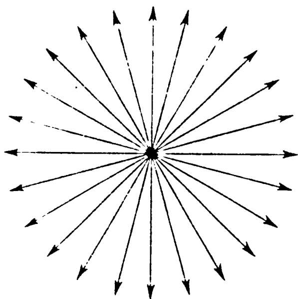
Łatwo stwierdzić, że jeśli znamy kierunki dwóch wektorów, np. **E** i **k** lub **H** i **k**, to kierunek trzeciego wektora jest określony jednoznacznie. Dlatego też w dalszych naszych rozważaniach w tym rozdziale będziemy mówili tylko o wektorze kierunkowym **k** i o wektorze natężenia pola elektrycznego **E**. Wektor **E** nazywany jest często wektorem świetlnym.

Zależność E od czasu w określonym punkcie przestrzeni, leżącym na promieniu fali, odpowiada zależności

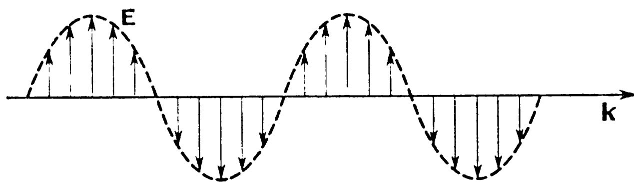
$$E = E_0 \sin(\omega t + \varphi),$$

gdzie E oznacza wartość natężenia pola elektrycznego w chwili t, E_0 — amplitudę E, ω — pulsację E (zwaną też częstotliwością kątową lub kołową), $(\omega t + \varphi)$ — fazę.

Rozważmy światło rozchodzące się bezpośrednio od źródła (tzw. *światło naturalne*). Choćby źródło miało najmniejsze rozmiary, jest w nim bardzo wiele atomów odpowiedzialnych za wysyłanie fal elektromagnetycznych. Dzięki tej mnogości atomowych źródeł w promieniowaniu, biegnącym w kierunku **k** bezpośrednio od źródła, występują zmiany wektora świetlnego we wszystkich kierunkach prostopadłych do wektora **k**. Takie *promieniowanie* nazywamy *niespolaryzowanym* (rys. 26.1). (Wektor **k** jest prostopadły do płaszczyzny rysunku.)



Rys. 26.1



Rys. 26.2

Jeżeli potrafimy zmiany wektora **E** we wszystkich falach składowych sprowadzić do jednej płaszczyzny, zawierającej wektor **k**, mamy do czynienia ze *światłem liniowo spolaryzowanym* (rys. 26.2). Płaszczyzna przechodząca przez wypadkowy wektor **E** i **k** nosi nazwę *płaszczyzny drgań wektora świetlnego*. Płaszczyznę do niej prostopadłą (zawierającą wektor **H** i **k**) nazywamy — według umowy przyjętej w optyce — *płaszczyzną polaryzacji*.

W przypadku fali spolaryzowanej liniowo zmiany **E** odbywają się stale w tej samej płaszczyźnie zawierającej wektor **k**. Sposoby wywoływania polaryzacji liniowej omówimy w następnych punktach.

Oprócz światła niespolaryzowanego i liniowo spolaryzowanego rozróżniamy jeszcze światło częściowo spolaryzowane, spolaryzowane kołowo i eliptycznie. W świetle spolaryzowanym częściowo mamy wiele płaszczyzn drgań wektora elektrycznego; amplitudy drgań wektora **E** nie są jednak jednakowe (rys. 26.3). Kierunek wektora **k** jest prostopadły do płaszczyzny rysunku.

Światło spolaryzowane kołowo może powstać przy nakładaniu się dwóch promieniowań spolaryzowanych liniowo. Wniosek taki można wyciągnąć korzystając z przedstawionych w § 10.8 rozważań, dotyczących składania dwóch ruchów harmonicznych wzajemnie do siebie prostopadłych. Wiemy już, że składając dwa drgania harmoniczne o jednakowych amplitudach i okresach, odbywające się w kierunkach wzajemnie prostopadłych i różniące się w fazie o $\pi/2$ (tzn. gdy w jednym ruchu mamy przejście przez położenie równowagi, to w drugim mamy maksymalną wartość wychylenia), otrzymuje się ruch jednostajny po kole.

Skoro zmiany wektora **E** w fali świetlnej odbywają się zgodnie z równaniem

$$E = E_0 \sin(\omega t + \varphi),$$

to są one analogiczne do zmian wychylenia od położenia równowagi w ruchu harmonicznym:

$$s = A\sin(\omega t+\varphi).$$

A zatem dwa światła spolaryzowane liniowo, wzajemnie prostopadłe, o tej samej amplitudzie E_0, lecz różniące się w fazie o $\pi/2$, np.

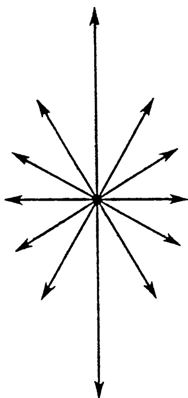
$$E_x = E_0\sin\omega t,$$

$$E_y = E_0\sin\left(\omega t+\frac{\pi}{2}\right) = E_0\cos\omega t,$$

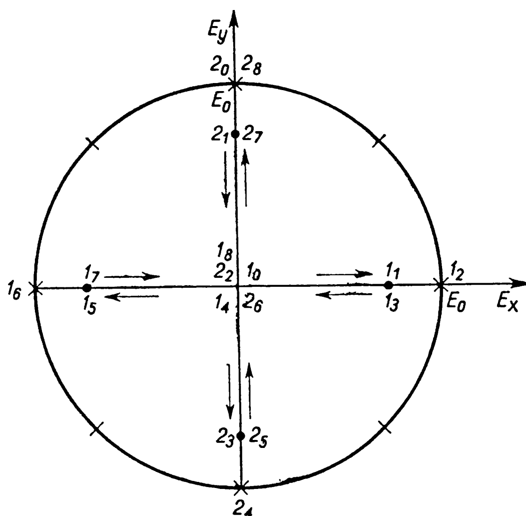
dadzą przy nałożeniu światło spolaryzowane kołowo. Wektor wypadkowy **E** będzie miał stałą długość

$$E = \sqrt{E_x^2+E_y^2} = E_0,$$

lecz będzie zmieniał swój kierunek w płaszczyźnie prostopadłej do **k**, obiegając jeden raz koło ruchem jednostajnym w ciągu jednego okresu.



Rys. 26.3



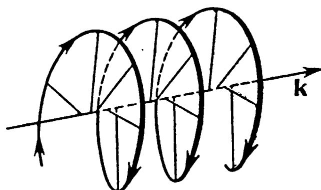
Rys. 26.4

Na rysunku 26.4 symbole 1_0, 1_1, 1_2, 1_3 itd. oznaczają kolejne wartości E_x co $\frac{1}{8}$ okresu. Analogicznie oznaczenia 2_0, 2_1, 2_2, 2_3 itd. odpowiadają kolejnym wartościom E_y co $\frac{1}{8}$ okresu. W wyniku sumowania obu drgań otrzymujemy okrąg o promieniu E_0. Uwzględniając dodatkowo rozchodzenie się fali świetlnej w kierunku promienia, łatwo stwierdzimy, że w przypadku fali świetlnej spolaryzowanej kołowo koniec wektora świetlnego opisuje linię śrubową (rys. 26.5).

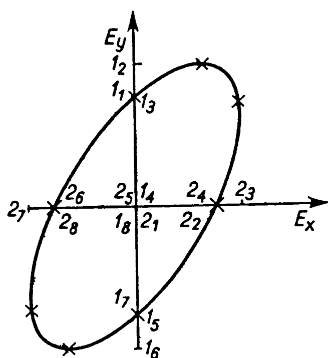
Gdy różnica faz obu składowych nie równa się zeru ani $\pi/2$, otrzymujemy w wyniku złożenia obu fal spolaryzowanych liniowo promieniowanie spolaryzowane eliptycznie. W tym przypadku nie jest konieczne, aby amplitudy były jednakowe i kierunki drgań obu wektorów składowych wzajemnie prostopadłe (rys. 26.6). Podany wynik opieramy

znów na analogii do składania ruchów harmonicznych. Podobnie, składanie dwóch drgań spolaryzowanych liniowo o dowolnych amplitudach i dowolnych kierunkach daje światło spolaryzowane liniowo, jeśli nie ma początkowej różnicy faz (rys. 26.7). Inna jest jednak płaszczyzna drgań wektora świetlnego w drganiu wypadkowym.

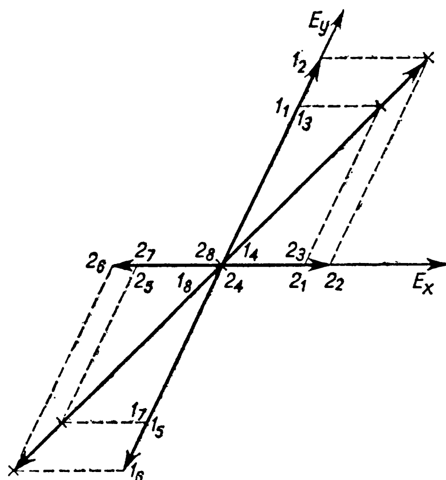
Każde drganie **E** w promieniowaniu niespolaryzowanym można przedstawić jako sumę geometryczną dwóch drgań wzajemnie prostopadłych. Dlatego też promieniowanie niespolaryzowane można przedstawić symbolicznie tak jak na rys. 26.8. Kropka na tym rysunku oznacza drganie prostopadłe do płaszczyzny rysunku.



Rys. 26.5



Rys. 26.6



Rys. 26.7



Rys. 26.8

Po tym ogólnym wstępie przechodzimy do omówienia kilku sposobów wywoływania liniowej polaryzacji światła. Rozróżnimy: polaryzację przy odbiciu od dielektryków przezroczystych, polaryzację przy załamaniu w ciałach izotropowych optycznie oraz polaryzację przy tzw. podwójnym załamaniu (w ośrodkach anizotropowych optycznie).

26.2. Polaryzacja światła przez odbicie

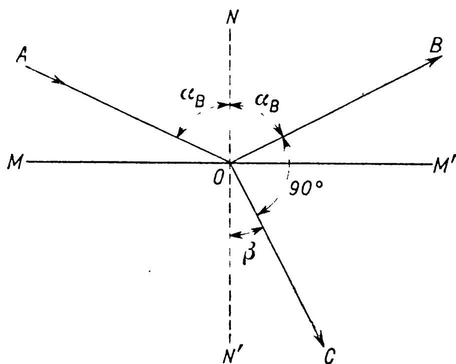
Liniową (częściową lub całkowitą) polaryzację światła można uzyskać stosując odbicie od przezroczystych dielektryków.

Niech promień *AO* pada na powierzchnię *MM'*, rozgraniczającą dwa ośrodki

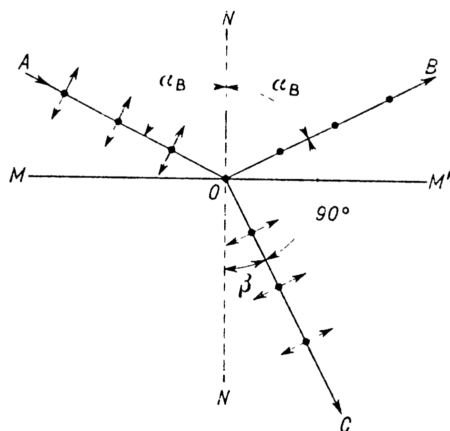
przezroczyste, np. powietrze i szkło (rys. 26.9). Promień OB jest promieniem odbitym, promień OC — promieniem załamanym. W tych warunkach, jak wykazuje doświadczenie, zarówno promień odbity, jak i załamany są częściowo spolaryzowane. Przy zmianie kąta padania zmienia się stopień uporządkowania drgań wektora **E** w świetle odbitym i w świetle załamanym. Brewster wykazał, że *jeśli promień odbity i załamany tworzą kąt 90°, to promień odbity jest całkowicie liniowo spolaryzowany*. Stosując prawa odbicia i załamania światła można prawo Brewstera podać w innej postaci:

$$n = \frac{\sin\alpha_B}{\sin\beta} = \frac{\sin\alpha_B}{\sin(90-\alpha_B)} = \operatorname{tg}\alpha_B. \tag{26.1}$$

Innymi słowy, *całkowita polaryzacja podczas odbicia występuje przy takim kącie padania którego tangens równa się współczynnikowi załamania*. Taki kąt padania nazywamy *kątem Brewstera* (α_B).



Rys. 26.9



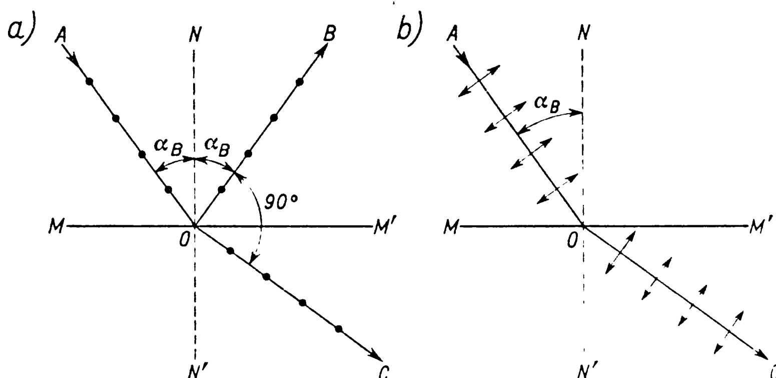
Rys. 26.10

Zagadnienia polaryzacji przy odbiciu nie rozpatrzyliśmy wyczerpująco, gdyż nie określiliśmy płaszczyzny drgań wektora świetlnego przy całkowitej liniowej polaryzacji. Przez dłuższy czas było nie wyjaśnione czy drgania wektora świetlnego odbywają się w płaszczyźnie padania, czy też w płaszczyźnie do niej prostopadłej. Sprawa ta rozstrzygnięta została dopiero po zbadaniu interferencji wiązek spolaryzowanych. Okazało się, że słuszne jest przedstawienie stanu polaryzacji tak, jak na rys. 26.10. Promień padający jest niespolaryzowany. Drgania wektora świetlnego w promieniu odbitym odbywają się w płaszczyźnie prostopadłej do płaszczyzny padania. Odbicie promienia niespolaryzowanego od płytki występuje zawsze niezależnie od jej ustawienia. Płytkę można obracać dokoła promienia padającego jako osi zachowując kąt padania Brewstera, nie powodując zaniku promienia odbitego.

Dotychczas mówiliśmy o odbiciu promieni niespolaryzowanych od ośrodków przezroczystych. Ciało odbijające odgrywało w tych warunkach rolę polaryzatora. Obecnie rozpatrzymy zagadnienie inne, a mianowicie odbicie promienia spolaryzowanego od powierzchni ośrodka przezroczystego. W ten sposób przekonamy się, że światło spo-

laryzowane ma inne właściwości niż światło niespolaryzowane, a mianowicie nie wykazuje symetrii względem kierunku rozchodzenia się światła.

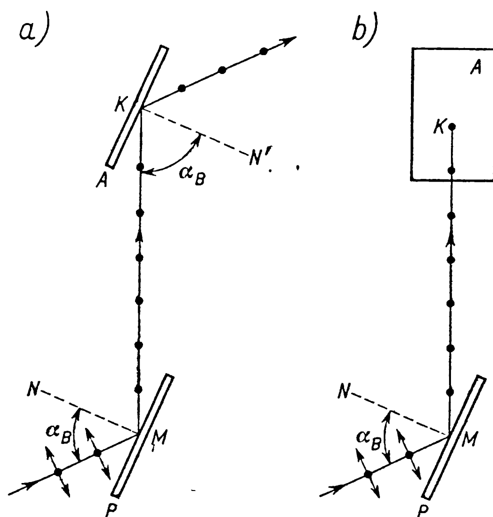
Niech promień spolaryzowany (np. przez odbicie od nie zaznaczonej na rys. 26.11 płytki szklanej) pada pod kątem Brewstera na powierzchnię MM' ośrodka przezroczystego. Jeśli drgania wektora świetlnego w promieniu padającym odbywają się w płaszczyźnie prostopadłej do płaszczyzny padania, to otrzymamy dwa promienie spolaryzowane: promień odbity i promień załamany (rys. 26.11a). Dobierzmy jednak tak promień padający lub obróćmy powierzchnię odbijającą tak, aby drgania wektora świetlnego w promieniu padającym odbywały się w płaszczyźnie padania. Otrzymamy wtedy tylko załamanie promienia (rys. 26.11b). Odbicia nie będzie. Odbiciu ulega ta składowa promieniowania padającego, w której drgania wektora świetlnego są prostopadłe do płaszczyzny padania, a padające promieniowanie spolaryzowane takiej składowej w ogóle nie zawiera. Łatwo się zorientować, że fakty te można wykorzystać do analizy, czy promień padający na płytkę jest spolaryzowany, czy też nie. Jeżeli promień padający jest niespolaryzowany, odbicie następuje przy każdym położeniu płaszczyzny padania (zmieniającym się podczas obrotu płytki dokoła padającego promienia, stanowiącego oś obrotu). Inaczej sprawa się przedstawia, gdy promień padający jest całkowicie spolaryzowany liniowo. Wtedy przy obrocie płytki dokoła promienia padającego jako osi z zachowaniem kąta padania Brewstera znajdziemy takie położenie, przy którym odbicia zupełnie nie będzie. Przy dalszym obrocie natężenie promienia odbitego stopniowo rośnie, osiąga maksimum po obrocie o 90°, a po obrocie o dalsze 90° znów spada do zera. Taka płytka odbijająca, sporządzona z przezroczystego materiału, obracana dokoła promienia padającego jako osi, działa jako analizator promieniowania.



Rys. 26.11

Na rysunku 26.12a i b pokazane jest zestawienie dwóch płytek szklanych użytych jako polaryzator P i analizator A. Odbicie od powierzchni polaryzatora P wywołuje całkowitą liniową polaryzację promienia biegnącego bezpośrednio od źródła, a więc niespolaryzowanego. (Dla zachowania przejrzystości na rys. 26.12a i b nie zaznaczono promieni załamanych.) Analizator A może być obracany dokoła promienia MK jako

osi, z zachowaniem kąta padania Brewstera. Na rysunku 26.12a płaszczyzny padania polaryzatora i analizatora są do siebie równoległe. A zatem drgania wektora **E** w promieniu odbitym od pierwszej płytki odbywają się w płaszczyźnie prostopadłej do płaszczyzny padania drugiej płytki. Promień zostaje częściowo odbity, a częściowo przechodzi przez drugą płytkę. Na rysunku 26.12b analizator jest tak ustawiony, że jego płaszczyzna padania jest prostopadła do płaszczyzny padania polaryzatora. Innymi słowy, drgania wektora **E** w promieniu odbitym od polaryzatora odbywają się w płaszczyźnie padania analizatora. Promień o takich drganiach nie ulega odbiciu, a tylko przechodzi przez analizator jako promień załamany.



Rys. 26.12

Na zakończenie powróćmy jeszcze do przypadku odbicia pod kątem różnym od kąta Brewstera. Pamiętamy, że powstają wtedy dwa promienie: odbity i załamany, oba częściowo spolaryzowane — pierwszy z przewagą składowej o kierunku drgań prostopadłym do płaszczyzny, drugi — z przewagą składowej o kierunku drgań równoległym do płaszczyzny padania. Stosunki natężeń promieniowania wiązki padającej I_p, wiązki odbitej I_o i wiązki załamanej I_z w zależności od kąta padania α i kąta załamania β wyznaczył Fresnel stosując teorię sprężystego eteru. Te same związki można wyprowadzić stosując założenia teorii elektromagnetycznej światła.

Dla drgań prostopadłych do płaszczyzny padania obowiązuje zależność

$$\frac{I_{o\perp}}{I_{p\perp}} = \frac{\sin^2(\alpha-\beta)}{\sin^2(\alpha+\beta)}, \tag{26.2a}$$

a dla drgań w płaszczyźnie padania —

$$\frac{I_{o\parallel}}{I_{p\parallel}} = \frac{\mathrm{tg}^2(\alpha-\beta)}{\mathrm{tg}^2(\alpha+\beta)}. \tag{26.2b}$$

W przypadku prawie prostopadłego padania promieni, gdy kąty α i β są bardzo małe, sinusy i tangensy można zastąpić samymi kątami. Stąd wniosek, że w tych warunkach

$$\frac{I_{o\perp}}{I_{p\perp}}=\frac{I_{o\parallel}}{I_{p\parallel}}=\frac{(\alpha-\beta)^2}{(\alpha+\beta)^2}$$

lub inaczej, po uwzględnieniu przybliżonej równości,

$$\frac{\sin\alpha}{\sin\beta}\approx\frac{\alpha}{\beta}=n,$$

$$\frac{I_{o\perp}}{I_{p\perp}}=\frac{I_{o\parallel}}{I_{p\parallel}}=\frac{(n-1)^2}{(n+1)^2}, \tag{26.3}$$

czyli w przypadku prawie prostopadłego padania promieni ułamek natężenia promieniowania ulegającego odbiciu jest jednakowy dla składowych o drganiach prostopadłych i równoległych i zależy tylko od współczynnika załamania n.

Na podstawie wzoru (26.3) można łatwo obliczyć, jaki procent ogólnego natężenia promieniowania padającego prawie prostopadle stanowi natężenie światła odbitego od powierzchni płytki szklanej o współczynniku załamania $n = 1{,}5$. A mianowicie:

$$\frac{I_o}{I_p}=\frac{I_{o\perp}+I_{o\parallel}}{I_{p\perp}+I_{p\parallel}}=\frac{(n-1)^2}{(n+1)^2}=\frac{0{,}5^2}{2{,}5^2}=\frac{1}{25},$$

czyli zaledwie 4%.

Zadania. 1. Jaki procent całkowitego natężenia niespolaryzowanego promieniowania padającego ulegnie odbiciu przy padaniu na płytkę szklaną ($n = 1{,}5$) pod kątem Brewstera?

2. Biorąc za podstawę (26.2b) wykazać, że: a) w przypadku kąta padania równego kątowi Brewstera światło odbite nie zawiera składowej o drganiach równoległych do płaszczyzny padania, b) w przypadku dowolnego kąta padania $\frac{I_{o\parallel}}{I_{p\parallel}}\leqslant\frac{I_{o\perp}}{I_{p\perp}}$.

26.3. Polaryzacja przy załamaniu światła w ośrodkach izotropowych

Od wzorów (26.2a i b) można łatwo (opierając się na zasadzie zachowania energii) przejść do wzorów określających stosunki natężeń w wiązce załamanej:

$$\frac{I_{z\perp}}{I_{p\perp}}=1-\frac{I_{o\perp}}{I_{p\perp}}, \tag{26.4}$$

$$\frac{I_{z\parallel}}{I_{p\parallel}}=1-\frac{I_{o\parallel}}{I_{p\parallel}}. \tag{26.5}$$

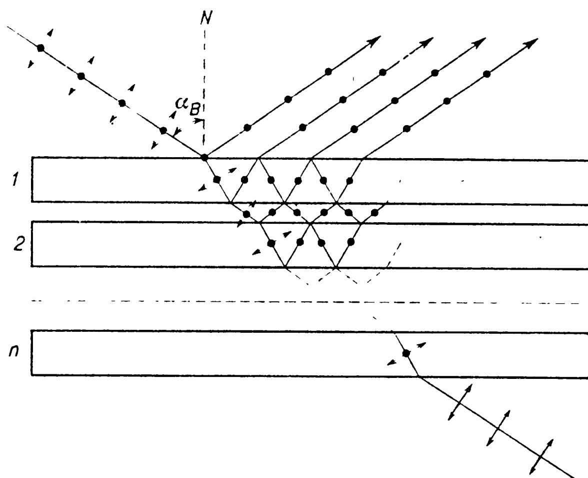
Z tych ostatnich wzorów i z wzorów (26.2a i b) wynika, że przy każdym kącie padania światła naturalnego (czyli niespolaryzowanego), nie wyłączając kąta padania Brewstera, promień załamany jest zawsze częściowo spolaryzowany.

Stosując wprowadzone ostatnio oznaczenia zdefiniujemy stopień polaryzacji P światła częściowo spolaryzowanego jako

$$P=\frac{I_{\parallel}-I_{\perp}}{I_{\parallel}+I_{\perp}},$$

czyli jako stosunek różnicy natężeń składowych o drganiach równoległych i prostopadłych do płaszczyzny padania do całkowitego natężenia.

Najwyższy stopień polaryzacji przy załamaniu na pojedynczej płytce szklanej o współczynniku załamania $n = 1,5$, w przypadku padającego pod kątem Brewstera światła naturalnego, wynosi zaledwie około 8% (por. zad. 1 niniejszego paragrafu). Jeśli taki promień załamany, częściowo spolaryzowany, rzucimy po wyjściu z płytki na drugą analogiczną płytkę płasko-równoległą pod kątem Brewstera (rys. 26.13), to znów w pro-



Rys. 26.13

mieniu odbitym znajdzie się pewien procent promieniowania o drganiach prostopadłych do płaszczyzny padania. Nowy promień załamany (przepuszczony) będzie miał wobec tego wyższy stopień polaryzacji. Przepuszczając światło przez szereg równoległych płytek szklanych zwiększamy stopień polaryzacji. Praktycznie biorąc, stos kilkunastu (dowolnie cienkich) płytek szklanych, na które rzucono światło naturalne pod kątem Brewstera, przepuszcza światło całkowicie liniowo spolaryzowane o drganiach wektora **E** w płaszczyźnie padania. Natężenie tego światła jest znacznie większe od natężenia światła odbitego, całkowicie spolaryzowanego, o drganiach prostopadłych do płaszczyzny padania.

Z a d a n i e. Wykorzystując wynik zad. 1 z § 26.2 obliczyć stopień polaryzacji promienia załamanego w płytce szklanej ($n = 1,5$) w przypadku padania światła niespolaryzowanego pod kątem Brewstera.

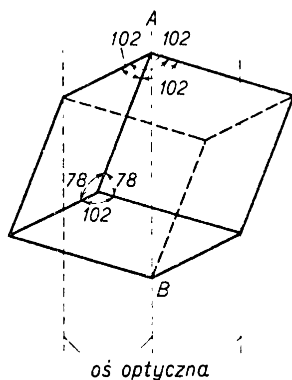
P y t a n i e. Scharakteryzować polaryzację promienia przepuszczonego przez płytkę płasko-równoległą, gdy światło padające pod kątem Brewstera jest całkowicie liniowo spolaryzowane i drgania wektora **E** odbywają się: a) w płaszczyźnie padania, b) w płaszczyźnie prostopadłej do płaszczyzny padania.

26.4. Polaryzacja przy podwójnym załamaniu światła

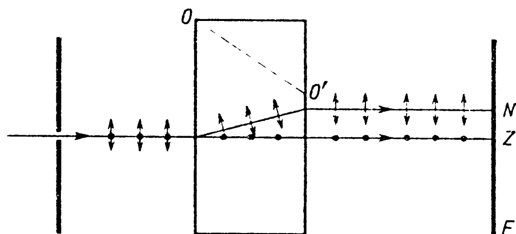
Omówienie trzeciego sposobu wywoływania polaryzacji światła musimy rozpocząć od krótkiej charakterystyki zjawiska podwójnego załamania, występującego w ośrodkach anizotropowych optycznie.

Znamy prawa rządzące załamaniem światła w ośrodkach jednorodnych izotropowych. Z wąskiej wiązki promieniowania powstaje pojedyncza wiązka załamana, do której stosują się oba podstawowe prawa załamania (§ 24.7).

Inaczej sprawa się przedstawia przy przejściu promieni przez ośrodki anizotropowe optycznie, np. przez kryształy (z wyjątkiem kryształów należących do układu regularnego, jak np. sól kuchenna NaCl). W tych warunkach powstają na ogół dwie wiązki załamane. Stąd nazwa — zjawisko podwójnego załamania. Rozpatrzmy to zjawisko na przykładzie przejścia promieni przez kryształy szpatu islandzkiego (kalcytu), czyli związku o składzie chemicznym $CaCO_3$. Wykorzystując płaszczyzny naturalnej łupliwości można uzyskać bryłkę szpatu o kształcie widocznym na rys. 26.14.



Rys. 26.14



Rys. 26.15

Doświadczenia można przeprowadzać rzucając promienie prostopadle na płytki prostopadłościenne (odpowiednio oszlifowane) rozmaicie wycięte względem osi symetrii kryształu. Wynik załamania przy przejściu przez takie płytki krystaliczne badamy na ekranie ustawionym poza płytką. Okazuje się, że w przypadku gdy płytka jest wycięta prostopadle do osi łączącej tępe naroża kryształu (*AB* na rys. 26.14), promień padający prostopadle na powierzchnię płytki nie ulega podwójnemu załamaniu. Na ekranie otrzymujemy pojedynczą plamkę świetlną. Oś *AB** jest osią główną kryształu. Każdy kierunek równoległy do *AB* stanowi kierunek tzw. *osi optycznej*, tzn. taki kierunek, w którym promienie przechodzą przez kryształ bez podwójnego załamania.

Szpat islandzki i inne kryształy, w których istnieje tylko jeden kierunek biegu promienia bez podwójnego załamania, nazywamy *kryształami jednoosiowymi*.

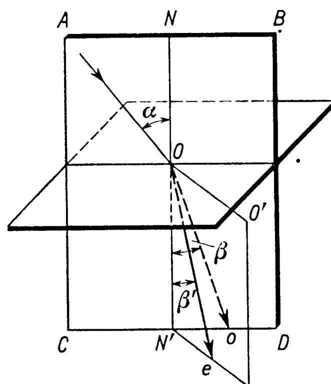
Rzućmy promień światła naturalnego prostopadle na powierzchnię płytki płasko-równoległej, wyciętej z kryształu jednoosiowego. Niech *OO'* będzie kierunkiem osi optycznej kryształu (rys. 26.15). Wtedy w płaszczyźnie rysunku otrzymamy w płytce dwa promienie: jeden — przechodzący bez załamania wzdłuż przedłużenia biegu promienia padającego, i drugi — promień załamany, wychodzący z płytki równolegle do promienia pierwszego. Oba promienie dadzą na ustawionym dalej ekranie *E* dwie

* *Oś główną kryształu szpatu* można zdefiniować bardziej ogólnie jako prostą przechodzącą przez tępe naroża i tworzącą jednakowe kąty ze wszystkimi spotykającymi się w tym narożu ścianami.

jasne plamki. Obracajmy płytkę dokoła promienia padającego jako osi. Okazuje się, że pierwszy ślad na ekranie, a mianowicie plamka Z, nie zmienia swego położenia, podczas gdy plamka N okrąża plamkę Z zakreślając koło. Promień wytwarzający plamkę Z nosi nazwę *promienia zwyczajnego*. Stosuje się on do praw załamania, które obowiązują w ośrodkach izotropowych, a więc padając prostopadle na kryształ wnika w niego bez załamania; przy innym kącie padania α załamuje się pod kątem β, przy czym spełniona jest zależność

$$\frac{\sin\alpha}{\sin\beta} = n_o = \frac{c}{v_o},$$

gdzie n_o jest wielkością stałą dla danego rodzaju promieniowania, niezależną od kierunku padania promienia, i stanowi *współczynnik załamania promienia zwyczajnego*. Ogólnie przyjęte jest stosowanie dla promienia zwyczajnego indeksu literowego „o" od nazwy francuskiej tego promienia — *ordinaire*. Tak więc n_o i v_o są to odpowiednio: współczynnik załamania i prędkość promienia zwyczajnego. Promień zwyczajny leży w płaszczyźnie padania.



Rys. 26.16

Drugi *promień*, zwany *nadzwyczajnym* (po francusku — *extraordinaire*; indeks literowy „e"), dający plamkę N, ma inne właściwości: 1) nie leży na ogół w płaszczyźnie padania, 2) stosunek $\sin\alpha$ do $\sin\beta$ nie ma wartości stałej i nie jest przyjmowany jako określenie współczynnika załamania. Promień nadzwyczajny leży w płaszczyźnie wyznaczonej przez kierunek osi optycznej i normalną padania. Tym się tłumaczy przesuwanie się plamki pochodzącej od promienia nadzwyczajnego przy obrocie płytki dokoła promienia padającego jako osi. Zmienia się przy tym w sposób ciągły położenie płaszczyzny wyznaczonej przez kierunek osi optycznej i normalną padania, a tym samym zmienia się położenie promienia nadzwyczajnego.

Na rysunku 26.16 przedstawiony jest bieg obu promieni przy kącie padania α różnym od zera. Rysunek ilustruje jednocześnie wyjście promienia nadzwyczajnego poza płaszczyznę padania $ABCD$.

Współczynnik załamania promienia nadzwyczajnego określany jest jako stosunek

prędkości światła w próżni do prędkości rozchodzenia się promienia nadzwyczajnego w danym kierunku w krysztale, a nie jako stosunek sinusów kąta padania i załamania. Prędkość promienia nadzwyczajnego równa się prędkości promienia zwyczajnego wtedy, gdy oba promienie biegną równolegle do osi optycznej. Dla każdego innego kierunku w krysztale prędkości te są różne. Prędkość promienia nadzwyczajnego (a zatem i jego współczynnik załamania) jest funkcją kąta, jaki tworzy promień z osią optyczną kryształu. Największa różnica prędkości, a więc i największa różnica współczynników załamania promienia nadzwyczajnego i zwyczajnego, występuje dla kierunków prostopadłych do osi optycznej. W tabelach współczynników załamania podawana jest zazwyczaj dla promienia nadzwyczajnego wartość n_e, odpowiadająca właśnie kierunkowi prostopadłemu do osi optycznej.

Tabela 26.1

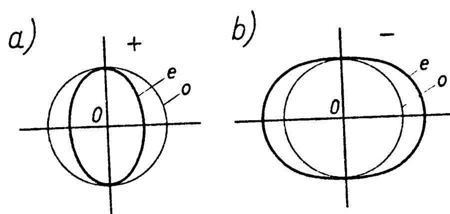
Współczynniki załamania kryształów jednoosiowych

Nazwa kryształu	Współczynnik załamania promienia zwyczajnego n_o	Współczynnik załamania promienia nadzwyczajnego n_e
apatyt	1,639	1,634
cynober	2,854	3,201
kwarc	1,544	1,553
lód	1,309	1,310
szpat islandzki	1,658	1,484
turmalin	1,642	1,622

W tabeli 26.1 mamy zestawienie wartości n_o i n_e, czyli tzw. głównych współczynników załamania dla kilku kryształów jednoosiowych (w odniesieniu do żółtej linii sodu). Jak widać z tabeli, dla niektórych kryształów główny współczynnik załamania promienia nadzwyczajnego jest większy, a dla innych — mniejszy od współczynnika załamania promienia zwyczajnego. Różnicę $n_e - n_o$ przyjmujemy za miarę *dwójłomności kryształu*. Gdy ta różnica jest dodatnia, kryształ nazywamy *optycznie dodatnim* i odwrotnie. Tak więc szpat islandzki jest przykładem kryształu ujemnego, a kwarc — dodatniego.

Powierzchnia falowa promienia zwyczajnego w krysztale ma kształt kuli, ponieważ prędkość tego promienia nie zależy od kierunku jego rozchodzenia się w krysztale. Powierzchnia falowa promienia nadzwyczajnego jest elipsoidą obrotową, wpisaną w kulę lub na niej opisaną. Punkty styczności kuli i elipsoidy leżą na osi optycznej. Największa różnica promienia kuli i półosi elipsoidy odpowiada kierunkowi prostopadłemu do osi optycznej. Na rysunku 26.17a i b przedstawione są przekroje powierzchni falowych w przypadku kryształów jednoosiowych, optycznie dodatnich i ujemnych. Analizując te rysunki należy pamiętać o tym, że im większa jest prędkość danego promienia w rozważanym ośrodku lub rozważanym kierunku, tym mniejszy jest współczynnik załamania. Punkty O potraktowane są jako źródła fali zwyczajnej i nadzwyczajnej w krysztale.

Wróćmy jeszcze do rys. 26.15. Badanie promienia zwyczajnego i nadzwyczajnego za pomocą dowolnego analizatora (np. za pomocą płytki szklanej odbijającej światło pod kątem Brewstera) wykazuje, że oba promienie są całkowicie liniowo spolaryzowane. Płaszczyzny drgań wektora świetlnego w obu promieniach są wzajemnie do siebie prostopadłe. Jeśli przy pewnym położeniu analizatora promień zwyczajny jest najsilniej odbijany, to przy tym samym położeniu promień nadzwyczajny w ogóle się nie odbija. Bliższe badania wykazały, że drgania wektora świetlnego w promieniu zwyczajnym odbywają się prostopadle do płaszczyzny przekroju głównego, zawierającej oś optyczną i promień zwyczajny. Drgania wektora świetlnego w promieniu nadzwyczajnym odbywają się w tej właśnie płaszczyźnie. Ten układ drgań widoczny jest na rys. 26.15.



Rys. 26.17

Analiza rys. 26.18 może być pomocna w przyswajaniu wiadomości dotyczących podwójnego załamania w kryształach jednoosiowych. Widzimy tam płytki krystaliczne rozmaicie wycięte w stosunku do osi optycznej. W każdym przypadku na powierzchnię płytki pada fala płaska niespolaryzowana, przedstawiona za pomocą trzech promieni prostopadłych do powierzchni płytki. Powierzchnia płytki pokrywa się w pewnej chwili z powierzchnią falową fali padającej. W tym momencie punkty padania zaznaczonych trzech promieni stają się — zgodnie z zasadą Huygensa — źródłem fal elementarnych, odpowiadających promieniowi zwyczajnemu i nadzwyczajnemu. Na rysunku 26.18a kierunek rozchodzenia się fal jest zgodny z kierunkiem osi optycznej — obie fale w tym kierunku poruszają się z tą samą prędkością. Kolejne czoło fali będzie wspólne dla promieni *e* i *o*. Rysunki 26.18b i c dotyczą przypadku, gdy kierunek fali padającej jest prostopadły do kierunku osi optycznej. Tym razem fale *e* i *o* idą w tym samym kierunku, lecz z różnymi prędkościami. Kolejne czoła obu fal są względem siebie przesunięte w przestrzeni.

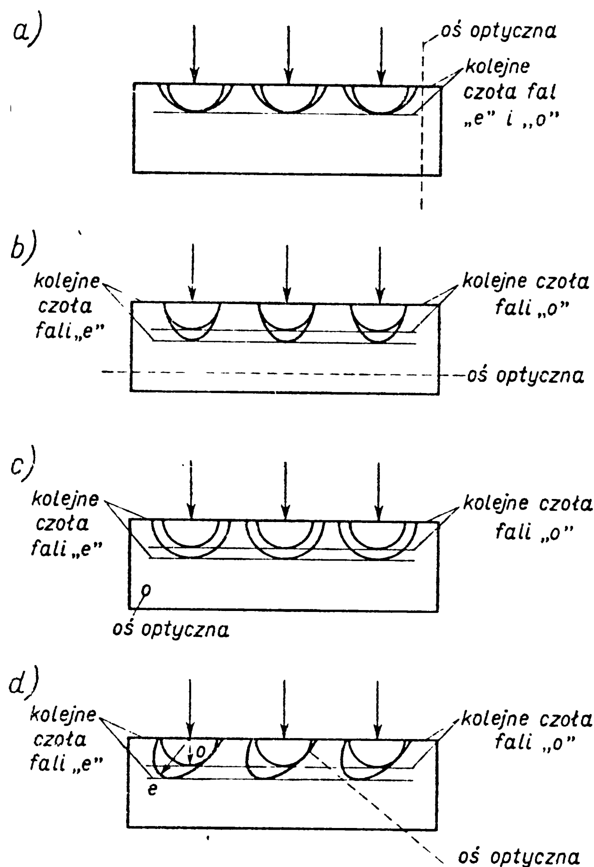
P y t a n i e. Dlaczego na tych dwóch rysunkach powierzchnie fal elementarnych *e* i *o* są różne?

Rysunek 26.18d przedstawia przypadek ogólniejszy: oś optyczna tworzy dowolny kąt z powierzchnią płytki, na którą w dalszym ciągu pada niespolaryzowana fala płaska. Tym razem nie tylko kolejne czoła fal *e* i *o* są rozsunięte przestrzennie, ale widać też wyraźnie rozsunięcie promieni *e* i *o*, czyli podwójne załamanie.

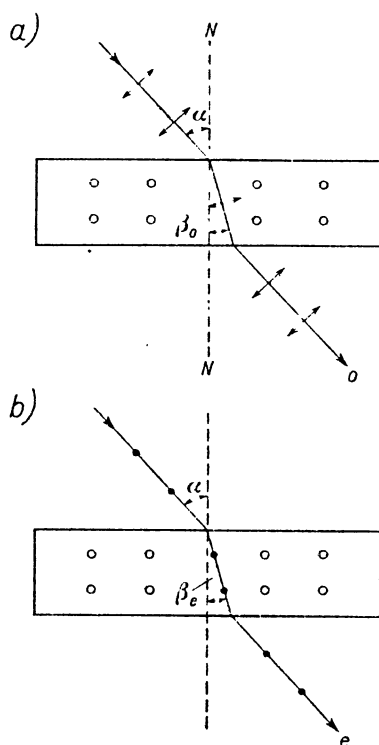
P y t a n i e. Czy można uważać, że rys. 26.18d i 26.15 dotyczą tego samego zjawiska?

Rozważane powyżej przykłady dotyczyły padania na płytkę światła niespolaryzowanego, z którego zawsze mogły powstać dwa promienie składowe o drganiach wektora

E wzajemnie prostopadłych. Rozpatrzmy z kolei przypadek padania na płytkę krystaliczną światła całkowicie liniowo spolaryzowanego. Na rysunku 26.19a przedstawione jest takie ustawienie płytki, że oś optyczna jest prostopadła do płaszczyzny rysunku (o czym przypominają małe kółka zaznaczone na rysunku), a drgania wektora **E** w promieniu padającym odbywają się w płaszczyźnie rysunku. W przypadku takiego doboru obu kierunków w płytce nie może powstać promień nadzwyczajny. Wychodzi z niej tylko promień zwyczajny, zgodnie z prawem Snelliusa. Gdy jednak kierunek drgań **E** jest równoległy do kierunku osi optycznej (rys. 26.19b), z płytki wychodzi tylko pro-



Rys. 26.18



Rys. 26.19

mień nadzwyczajny. Wyjątkowo w tym przypadku można wykorzystać do obliczenia kąta załamania promienia nadzwyczajnego wartość współczynnika załamania n_e podaną w tabeli i odwrotnie, z wyznaczonego doświadczalnie kąta załamania promienia nadzwyczajnego można obliczyć n_e z wzoru

$$\frac{\sin\alpha}{\sin\beta_e} = n_e.$$

P y t a n i e. Opisać zjawiska zachodzące w przypadku, gdy płytkę z rys. 26.19a obrócimy dokoła normalnej padania o kąt $0° < \gamma < 90°$ zachowując stan całkowitej liniowej polaryzacji promienia padającego.

Z a d a n i e. Wąska wiązka światła niespolaryzowanego pada na płytkę płasko-równoległą z kryształu szpatu islandzkiego pod kątem $\alpha = 45°$. Płaszczyzna padania jest prostopadła do osi optycznej płytki. Grubość płytki $d = 1$ cm. Zbadać bieg promieni e i o po wyjściu z płytki. Obliczyć ich wzajemne przesunięcie. Jakie efekty wywoła obrót polaryzatora wstawionego w bieg promienia padającego? (Osią obrotu jest kierunek promienia padającego.) Jakie rysunki stanowią ilustrację odpowiedzi na ostatnie pytanie?

U w a g a. W krysztale oba promienie biegną w płaszczyźnie prostopadłej do osi optycznej. W tych warunkach do obu promieni można stosować prawo Snelliusa, tzn.

$$\frac{\sin\alpha}{\sin\beta_o} = n_o, \qquad \frac{\sin\alpha}{\sin\beta_e} = n_e,$$

gdzie n_o i n_e są głównymi współczynnikami załamania szpatu, podanymi w tabeli.

Prócz kryształów jednoosiowych, należących do układu tetragonalnego i heksagonalnego, występują w przyrodzie także *kryształy dwuosiowe* (układ rombowy, jedno- i trójskośny). W takich kryształach istnieją dwa kierunki, w których promienie biegną z jednakowymi prędkościami. Oba promienie są jednak w tych kryształach „nadzwyczajne”, gdyż żaden z nich nie podlega prawom załamania obowiązującym w ośrodkach izotropowych.

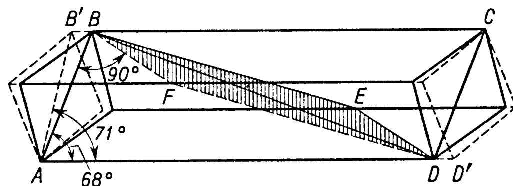
26.5. Polaryzatory i analizatory światła

Korzystając ze zjawiska podwójnego załamania światła w kryształach można budować dogodne polaryzatory i analizatory światła. W tym celu staramy się usunąć jeden z dwóch wychodzących z kryształu liniowo spolaryzowanych promieni. Można to uzyskać różnymi sposobami. Niektóre kryształy mają tę właściwość, że znacznie silniej pochłaniają jeden z promieni; wtedy praktycznie biorąc, przy dostatecznej grubości kryształu wychodzi z niego tylko jedna wiązka. Do takich *kryształów*, zwanych *dichroicznymi*, należy np. turmalin. Ujemną stroną turmalinu jest to, że wychodząca z niego wiązka nadzwyczajna e ma zabarwienie zielonobrunatne na skutek niejednakowego pochłaniania różnych składników światła białego, przechodzącego przez kryształ. Poza tym okazuje się, że płytki turmalinowe wycięte z różnych kryształów mają nieco różne właściwości optyczne. Zmniejsza to oczywiście bardzo możliwość ich praktycznego stosowania.

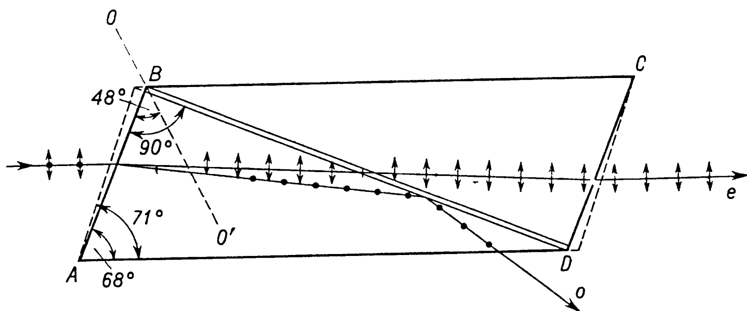
Bardzo pomysłowy sposób usunięcia jednego ze spolaryzowanych promieni zawdzięczamy Nicole'owi. Wykorzystał on kryształ szpatu islandzkiego, lecz wprowadził w nim następujące zmiany: zeszlifował powierzchnie ścian bocznych tak (rys. 26.20), że kąt między ściankami zmienił się z 71° na 68°. Następnie przeciął kryształ wzdłuż płaszczyzny przekątnej, również zeszlifowując powierzchnię tak, że kąt zaznaczony na rysunku wynosił 90°. Potem obie połówki kryształu skleił balsamem kanadyjskim. W ten sposób powstało urządzenie zwane *pryzmatem Nicole'a* lub krótko *nikolem*. Przekrój *ABCD* pryzmatu przedstawiony jest na rys. 26.21. Jest to równocześnie przekrój główny, zawiera bowiem kierunek osi optycznej.

Niech promień świetlny pada na powierzchnię boczną kryształu nikola równolegle

do jego podstawy. Przy wejściu do kryształu rozdziela się on na promień zwyczajny i nadzwyczajny. Promień nadzwyczajny biegnie dalej w tej samej płaszczyźnie (przekroju głównego, zawierającego oś optyczną i normalną padania), prawie nie odchylając się od pierwotnego biegu. Dochodzi on do cienkiej warstwy balsamu kanadyjskiego, przechodzi przez nią też prawie bez zmiany kierunku (cienka warstwa płasko-równoległa) i wychodzi dalej z kryształu, zachowując prawie nie zmieniony kierunek wiązki padającej.



Rys. 26.20



Rys. 26.21

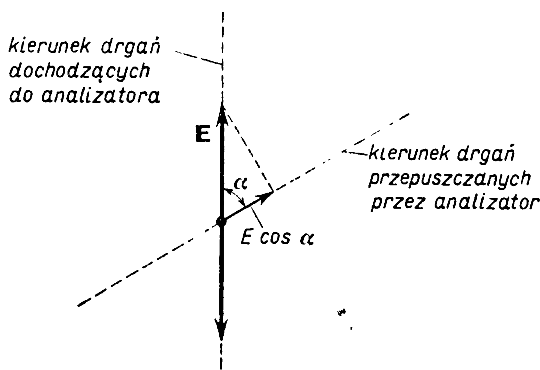
Promień zwyczajny zachowuje się inaczej. Wchodząc z powietrza do szpatu islandzkiego odchyla się przy załamaniu mocniej w stronę normalnej padania. Następnie pada na warstwę balsamu kanadyjskiego. Z porównania współczynników załamania promienia zwyczajnego dla szpatu (1,66) i dla balsamu (1,549) wynika, że szpat jest ośrodkiem optycznie gęstszym dla promienia zwyczajnego niż balsam kanadyjski. Kąt, pod jakim promień zwyczajny pada w nikolu na warstwę balsamu, jest większy od kąta granicznego. Spełnione są zatem warunki, w których występuje zjawisko całkowitego wewnętrznego odbicia. Promień zwyczajny nie wchodzi do balsamu, lecz dzięki całkowitemu wewnętrznemu odbiciu zostaje skierowany na zaczernioną podstawę nikola i tam zostaje pochłonięty.

Z nikola wychodzi zatem tylko promień nadzwyczajny. Kierunek jego jest prawie zgodny z kierunkiem promienia padającego. Drgania wektora świetlnego odbywają się w nim w płaszczyźnie przekroju głównego *ABCD*. Wynika to z rozważań ogólnych, dotyczących kierunków drgań wektora **E** w promieniach załamanych *o* i *e* (por. § 26,4, rys. 26.15). Prostopadle do płaszczyzny przekroju głównego *ABCD* odbywają się

drgania **E** w promieniu zwyczajnym. W płaszczyźnie tej natomiast odbywają się drgania **E** w promieniu nadzwyczajnym, wychodzącym z nikola. Geometrycznie płaszczyznę tę określają przekątne *AB* i *CD* rombowych podstaw nikola (rys. 26.20), łączące wierzchołki kątów rozwartych.

Pojedynczy pryzmat nikola spełnia rolę polaryzatora w stosunku do promieniowania idącego bezpośrednio od źródła światła. Ten sam pryzmat może też spełniać rolę analizatora. Jeżeli będziemy go obracali dokoła naturalnego promienia padającego jako osi, to nie zaobserwujemy żadnych zmian natężenia przechodzącego promieniowania. Jeżeli jednak zaczniemy obracać nikol dokoła całkowicie liniowo spolaryzowanego padającego promienia jako osi, to stwierdzimy zmiany w natężeniu przechodzącego przez nikol światła. Gdy płaszczyzna przekroju głównego *ABCD* nikola jest zgodna z płaszczyzną drgań padającego spolaryzowanego promienia, to przez nikol przechodzi światło o maksymalnym natężeniu (por. rys. 26.19b). Jeżeli skręcimy nikol o 90° do położenia takiego, przy którym jego płaszczyzna przekroju głównego jest prostopadła do płaszczyzny drgań padającego promieniowania, to nie będzie takiej składowej, która mogłaby przejść przez nikol (rys. 26.19a). Nastąpi całkowite wygaszenie światła. Dalszy obrót o 90° doprowadzi znów do maksimum natężenia światła. Całkowicie spolaryzowany promień padający na nikol-analizator można otrzymać z nikola-polaryzatora. Prawo Malusa określa natężenie I promieniowania wychodzącego z nikola-analizatora w zależności od kąta α, jaki tworzą ze sobą płaszczyzna drgań promienia padającego na analizator i płaszczyzna *ABCD* jego przekroju głównego:

$$I = I_0 \cos^2 \alpha, \tag{26.6}$$



Rys. 26.22

gdzie I_0 oznacza natężenie wiązki wychodzącej z nikola w przypadku równoległego ustawienia obu wymienionych płaszczyzn. Uzasadnienie tego wzoru wiąże się z dwoma faktami: 1) składowa drgań przechodzących przez analizator ma amplitudę $E \cos \alpha$ (rys. 26.22), gdy amplituda drgań padających na niego wynosi E, 2) natężenie jako ilość energii przechodzącej w jednostce czasu przez jednostkę powierzchni jest proporcjonalna do kwadratu amplitudy. Ze wzoru widzimy, że układ dwóch nikoli (polaryzatora i analizatora) pozwala na zmianę w sposób ciągły natężenia światła wychodzą-

cego. W tym celu zmieniamy kąt α przez obrót jednego z nikoli dokoła promienia padającego jako osi.

Zanim przejdziemy do omówienia praktycznego zastosowania nikoli w tzw. polarymetrach, wspomnimy krótko o osiągnięciach, jakie przyniosły ostatnie dziesiątki lat. Stosowanie pryzmatów nikoli w przyrządach optycznych natrafiało na duże trudności. Nikole otrzymywane są z naturalnych kryształów szpatu islandzkiego, a te mają niewielkie rozmiary i stosunkowo rzadko występują w przyrodzie. Mimo małych rozmiarów (stanowiących w wielu przypadkach wielką wadę) nikole są dość grube i w przyrządach optycznych zajmują stosunkowo dużo miejsca. Co więcej, wiązki padające na nikol muszą być wiązkami kierunkowymi: muszą to być wiązki wąskie, biegnące równolegle do podstawy nikola.

P y t a n i e. Wyjaśnić, dlaczego dla wiązek mocno odchylonych od kierunku równoległego do krawędzi *AD* nikol nie spełnia swego zadania?

Nic więc dziwnego, że prowadzone były wytężone prace, mające na celu wykonanie urządzeń polaryzacyjnych o szerszym zastosowaniu niż mogą mieć nikole.

W urządzeniach polaryzacyjnych, przewyższających zaletami nikole, stosuje się różne kryształy dichroiczne, m.in. kryształy syntetyczne kwaśnego trójjodosiarczanu chininy. Kryształy te zbadane zostały w roku 1852 przez lekarza angielskiego Herapatha, dlatego nazwano je kryształami herapatytu. Mają one postać drobnych igieł. Herapath usiłował otrzymać kryształki większych rozmiarów, ale mimo wielkiego wkładu pracy pomyślnych wyników nie uzyskał. Późniejsze prace rozwijały się w innym kierunku. Zrezygnowano z otrzymania pojedynczych dużych kryształów. Starano się otrzymać kryształy możliwie jak najmniejsze, ale jednakowo zorientowane, tzn. ustawione tak, aby osie poszczególnych igieł krystalicznych były do siebie równoległe. Ostatecznie otrzymano kryształki herapatytu o długości około mikrometra i o grubości rzędu ułamków mikrometra. W rozmaity sposób starano się uzyskać ich jednakowe ustawienie. Stosowano m.in. do tego celu działanie silnych pól magnetycznych i elektrycznych; były to jednak metody kosztowne, a nie dające zadowalających rezultatów.

Do otrzymywania tzw. płyt polaroidowych stosuje się ostatnio następującą metodę. Ciekłą masę plastyczną o odpowiedniej zawartości herapatytu przepuszcza się przez specjalne „sita", zawierające dużą liczbę wąskich otworów — szczelin. Podczas przeciskania masy plastycznej przez te szczeliny zawarte w niej kryształki herapatytu ustawiają się tak, że ich osi podłużne przyjmują położenia równoległe do siebie. Bezpośrednio po wyjściu z sita masa plastyczna zastyga zachowując uporządkowany stan kryształków herapatytu. Taką zastygniętą masę umieszcza się między płytami szklanymi w celu zabezpieczenia od mechanicznych uszkodzeń i otrzymuje się w ten sposób polaroid w ostatecznej postaci.

Uporządkowanie kryształków można też otrzymywać przez odpowiednie rozcieranie krzepnącej masy plastycznej z dichroicznymi mikrokryształami między płytkami szklanymi.

Z licznych zastosowań polaroidów wymienimy tu kilka. W wielu przypadkach spotykamy się z rażącymi odblaskami utrudniającymi oglądanie pewnych przedmiotów. Wystarczy przypomnieć odbicie od powierzchni morza lub jeziora, od obrazów olej-

nych na wystawie malarstwa, od szkła przykrywającego zdjęcie fotograficzne, od szyby wystawowej itp. Światło odbite jest zawsze co najmniej częściowo spolaryzowane. Okulary polaroidowe odpowiednio ustawione zmniejszają natężenie przechodzącego przez nie spolaryzowanego promieniowania, a zatem ułatwiają — w wymienionych wyżej przykładach — oglądanie przedmiotów na wystawach lub wypatrywanie obiektów na powierzchni morza.

Ponieważ światło rozproszone jest również częściowo spolaryzowane, więc płyty polaroidowe mogą spełniać rolę filtrów przy fotografowaniu, pozwalając na robienie bardziej kontrastowych zdjęć, np. chmur.

Poważnym odbiorcą płyt polaroidowych będzie prawdopodobnie przemysł samochodowy. Chodzi tu o usunięcie możliwości oślepiania kierowcy światłem reflektorów samochodu nadjeżdżającego z przeciwka. Wyobraźmy sobie, że zarówno osłony szklane reflektorów, jak i szyby przednie w samochodzie zastąpione zostały płytami polaroidowymi. Płyty te osadzone są w ten sposób, że zapewniają określony kierunek drgań (np. pod kątem 45° do poziomu od lewej strony do prawej w dół) przepuszczanego promieniowania. W przypadku jednakowego ustawienia płyt reflektorów i przedniej szyby, do oczu kierowcy dochodzi światło powstające wskutek odbicia lub rozproszenia promieniowania pochodzącego od reflektorów jego samochodu. Inaczej się sprawa przedstawia z promieniowaniem pochodzącym od reflektorów samochodu jadącego z przeciwka. Kierunek drgań przepuszczanych przez płyty polaroidowe nadjeżdżającego samochodu jest prostopadły do kierunku drgań przepuszczanych przez szyby samochodu pierwszego. A zatem światło reflektorów samochodu drugiego nie przejdzie przez polaroidową płytę samochodu pierwszego i nie oślepi kierowcy. Płyty polaroidowe reflektorów samochodu drugiego i płyta polaroidowa samochodu pierwszego stanowią odpowiednik skrzyżowanych nikoli, pochłaniających całkowicie światło. Kierowca samochodu pierwszego dojrzy nadjeżdżający samochód tylko dzięki światłu odbitemu, pochodzącemu od reflektorów własnego samochodu.

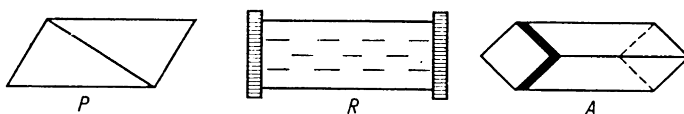
26.6. Skręcanie płaszczyzny polaryzacji

Dwa nikole lub dwa polaroidy oddają cenne usługi przy badaniu zjawiska skręcania płaszczyzny polaryzacji ściśle związanego ze strukturą ciał, przez które przechodzi światło liniowe spolaryzowane.

Zjawisko skręcania płaszczyzny polaryzacji odkryte zostało w roku 1811 przez Arago podczas badania podwójnego załamania w kwarcu krystalicznym. Do dnia dzisiejszego odkryto tysiące rozmaitych ciał, zarówno krystalicznych, jak bezpostaciowych lub np. w postaci roztworów, wykazujących mniejszą lub większą zdolność skręcania płaszczyzny polaryzacji. Stwierdzono między innymi, że właściwość tę ma wodny roztwór cukru i że można ją wykorzystać do dokładnego określania stężenia roztworu, co znalazło szerokie zastosowanie w przemyśle cukrowniczym. Omówimy dokładniej to zjawisko.

Wstawiając między skrzyżowane nikole P i A (rys. 26.23) rurkę szklaną R wypełnioną roztworem cukru uzyskujemy rozjaśnienie pola widzenia. Przez odpowiedni obrót

analizatora można jednak otrzymać całkowite wygaszenie światła. Stąd wniosek, że stan liniowej polaryzacji został zachowany, uległ jednak zmianie kierunek drgań wektora **E** — zmieniła się płaszczyzna drgań wektora **E** (zawierająca wektory **E** i **k**), jak również prostopadła do niej płaszczyzna polaryzacji. Skręcenie płaszczyzny polaryzacji wywołane jest działaniem cząsteczek cukru na drodze promieni świetlnych. Kąt skręcenia płaszczyzny polaryzacji ustalamy odczytując na podziałce kąt, o jaki należy obrócić analizator, aby po wstawieniu między skrzyżowane nikole rurki z roztworem cukru ponownie otrzymać zupełne zaciemnienie pola widzenia.



Rys. 26.23

Doświadczenia wykazały, że kąt skręcenia płaszczyzny polaryzacji α w danej temperaturze i dla danej długości fali jest proporcjonalny do długości l drogi promienia w roztworze (długość rurki) i do stężenia k roztworu

$$\alpha = \alpha_0 kl. \tag{26.7}$$

Współczynnik proporcjonalności α_0 wyraża według umowy kąt skręcenia płaszczyzny polaryzacji, występujący przy użyciu roztworu o stężeniu 1 kg/0,1 m^3, czyli 1 g/100 cm^3, gdy długość drogi promienia $l = 0{,}1$ m.

Stężenie roztworu wyznaczamy zazwyczaj mierząc kąty skręcenia α_1 i α_2 przy wypełnieniu tej samej rurki roztworem wzorcowym o stężeniu k_1 i roztworem badanym o stężeniu k_2. Ze wzoru

$$\frac{\alpha_2}{\alpha_1} = \frac{\alpha_0 k_2 l}{\alpha_0 k_1 l}$$

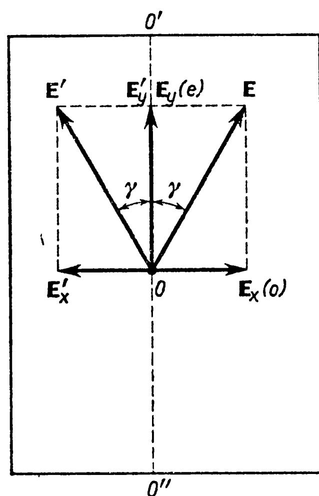
otrzymujemy

$$k_2 = k_1 \frac{\alpha_2}{\alpha_1}.$$

Dokładność pomiaru stężenia zależy od dokładności pomiarów obu kątów. Taki przebieg doświadczenia, jak omówiony wyżej, nie zapewnia dużej dokładności, gdyż oko z trudem ustala moment ustawienia analizatora na najgłębszą ciemność. Przy dokładnych pomiarach kąta skręcenia wykorzystuje się najczęściej zdolność oka do wykrywania różnic w oświetleniu stykających się powierzchni. Ta zdolność oka została m.in. wykorzystana w przyrządzie przeznaczonym do wyznaczania stężenia roztworu cukru, zwanym sacharymetrem półcieniowym Laurenta.

W sacharymetrze półcieniowym Laurenta, prócz zestawienia dwóch nikoli i rurki z roztworem cukru, mamy dodatkową płytkę kwarcową zwaną półfalówką, przesłaniającą połowę pola widzenia i przyczyniającą się do różnego oświetlenia obu jego części.

Półfalówka jest to płytka kwarcowa wycięta równolegle do osi optycznej $O'O''$ (rys. 26.24). Niech punkt O będzie śladem wiązki promieniowania liniowo spolaryzowanego, padającego prostopadle na powierzchnię płytki. Kierunek drgań wektora elektrycznego **E** w tym promieniu tworzy mały kąt γ z kierunkiem osi optycznej $O'O'$, płytki. Wektor ten rozkłada się w płytce na dwa: jeden, $\mathbf{E}_y$, równoległy do osi (promień nadzwyczajny), drugi, $\mathbf{E}_x$, prostopadły do osi (promień zwyczajny). Ponieważ kwarc należy do kryształów jednoosiowych dodatnich, prędkość promienia nadzwyczajnego jest mniejsza od prędkości promienia zwyczajnego. W związku z tym przy przechodzeniu obu promieni przez płytkę powstaje między nimi różnica faz. Grubość płytki jest tak dobrana, że wektor elektryczny drgań równoległych do osi (promień nadzwyczajny) po przejściu pızez płytkę wyprzedza w fazie o π wektor elektryczny drgań prostopadłych do osi. Odpowiada to różnicy dróg optycznych równej połowie długości fali — stąd nazwa półfalówka (por. przykład rozważany dalej). Po wyjściu z płytki oba promienie wchodzą do analizatora, który rozkłada każdy z nich na promień zwyczajny i nadzwyczajny. Poza analizator wychodzi tylko promień nadzwyczajny, który jest wynikiem złożenia obu składowych promieni nadzwyczajnych utworzonych w analizatorze. Składanie tych promieni jest możliwe dzięki temu, że obie składowe są spójne i drgania odbvwaią się w jednej płaszczyźnie.

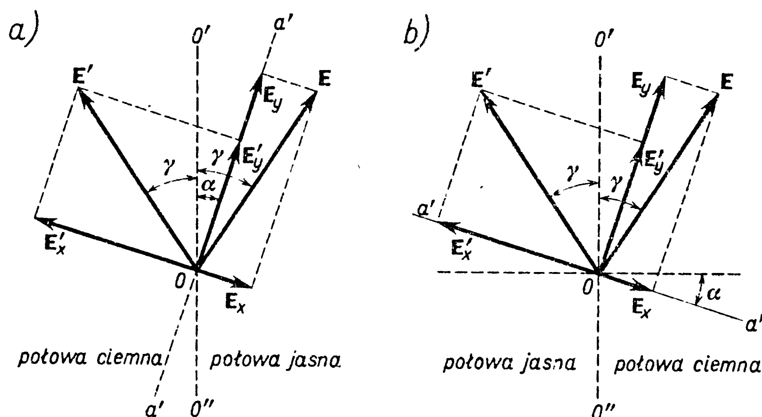


Rys. 26.24

Prześledzenie graficzne opisanego wyżej zachowania się wektorów świetlnych w płytce i w analizatorze byłoby skomplikowane. Postępujemy inaczej. Upraszczamy zagadnienie zakładając, że składanie drgań odbywa się po wyjściu z płytki przed wejściem do analizatora. Możemy sobie na to pozwolić, gdyż nie ma to wpływu na wynik ostateczny. Na rysunku 26.24 wektor **E** przedstawia wektor świetlny liniowo spolaryzowanego światła padającego na płytkę. Wektor ten rozkłada się w płytce na składowe $\mathbf{E}_x$ i $\mathbf{E}_y$.

Wskutek opóźnienia w fazie jednego z tych wektorów o π mamy w chwili wychodzenia promienia z płytki wektory $\mathbf{E}'_y = \mathbf{E}_y$ i $\mathbf{E}'_x = -\mathbf{E}_x$. W wyniku składania otrzymujemy wektor wypadkowy $\mathbf{E}'$, tzn. po przejściu przez płytkę zmieniła się płaszczyzna drgań wektora świetlnego. Nastąpiło skręcenie płaszczyzny drgań o kąt $= 2\gamma$.

Wyobraźmy sobie, że taką półfalówkę ustawiliśmy za polaryzatorem, zasłaniając nią połowę pola widzenia. Efekt jest taki, że do analizatora dochodzą w każdej połowie pola widzenia wektory świetlne o różnych kierunkach, o kierunku $\mathbf{E}$ (pierwotnym) i kierunku $\mathbf{E}'$ (obróconym działaniem płytki). Oba kierunki drgań są symetryczne względem kierunku osi optycznej $O'O''$. Gdy kierunek drgań przepuszczanych przez analizator jest równoległy do $O'O''$, analizator wygasza składowe $\mathbf{E}_x$ i $\mathbf{E}'_x$, a przepuszcza składowe $\mathbf{E}_y$ i $\mathbf{E}'_y$. Ponieważ składowe $\mathbf{E}_y$ i $\mathbf{E}'_y$ są równe w obu połówkach pola widzenia, są one jednakowo oświetlone. Gdy skręcimy analizator o kąt 90°, wygaszone zostaną składowe $\mathbf{E}_y$ i $\mathbf{E}'_y$, a przepuszczone zostaną składowe $\mathbf{E}_x$ i $\mathbf{E}'_x$. Ponieważ ich wartości liczbowe są też sobie równe, więc i w tym przypadku obie połówki pola widzenia są jednakowo, choć słabiej oświetlone.



Rys. 26.25

Położenia równej jasności obu przylegających do siebie połówek pola widzenia można ustalić z dużą dokładnością (do 0,01°) przy użyciu światła monochromatycznego. Spróbujmy bowiem sprawdzić, co się dzieje, gdy kierunek $a'a'$ drgań przepuszczanych przez analizator tworzy mały kąt α z kierunkiem osi optycznej płytki $O'O''$ (rys. 26.25a) lub z kierunkiem do niego prostopadłym (rys. 26.25b). W tych przypadkach składowe drgań przepuszczanych w obu połówkach pola widzenia mają nierówne amplitudy, a zatem gdy po osiągnięciu równomiernego oświetlenia całego pola widzenia dalej obracamy analizator, jedna z połówek pola widzenia rozjaśnia się, a druga — zaciemnia. Wprowadzenie półfalówki umożliwia dokładne określenie położenia odpowiadającego równym oświetleniom obu połówek pola widzenia. Po ustaleniu tego położenia wprowadzamy rurkę z roztworem cukru między płytkę i analizator. Wskutek skręcenia

płaszczyzny polaryzacji jednakowe oświetlenia uzyskamy ponownie po obrocie analizatora o kąt równy kątowi skręcenia płaszczyzny polaryzacji przez roztwór cukru.

Światło użyte do badań powinno być monochromatyczne. Można je otrzymać ze źródła światła monochromatycznego, np. z lampy sodowej lub też dzięki użyciu odpowiedniego filtru. Zazwyczaj stosowane jest sodowe światło żółte. Do jego długości fali dostosowana jest grubość płytki kwarcowej, zapewniającej opóźnienie w fazie promienia zwyczajnego o π.

P r z y k ł a d 1. Potraktujmy jako zadanie obliczenie grubości płytki kwarcowej (wyciętej równolegle do osi optycznej), która przy padaniu prostopadłym monochromatycznego światła spolaryzowanego zachowuje się jak półfalówka, czyli zapewnia różnicę faz wychodzących promieni równą π.

Tworzenie się różnicy faz wychodzących promieni w półfalówce (podobnie jak w omawianej w § 26.7 ćwierćfalówce) można łączyć ze zmianą długości fali λ_e i λ_o w stosunku do długości fali λ w próżni. Jak już wiemy

$$\lambda_e = \frac{\lambda}{n_e}, \qquad \lambda_o = \frac{\lambda}{n_o}.$$

Liczba długości fal mieszczących się w grubości d płytki wynosi odpowiednio

$$N_e = \frac{d}{\lambda_e} = \frac{dn_e}{\lambda}, \qquad N_o = \frac{d}{\lambda_o} = \frac{dn_o}{\lambda}.$$

W kryształach optycznie dodatnich, takich jak kwarc, $n_e > n_o$, a zatem $N_e > N_o$, czyli promień nadzwyczajny wyprzedza w fazie promień zwyczajny. W przypadku półfalówki

$$N_e - N_o = \tfrac{1}{2},$$

czyli

$$\frac{d(n_e - n_o)}{\lambda} = \tfrac{1}{2},$$

skąd łatwo można obliczyć grubość półfalówki przy znanych wartościach n_e, n_o i λ światła padającego. Na przykład dla światła sodowego o długości $\lambda = 5890$ Å grubość kwarcowej półfalówki wynosi 0,032 mm.

Na zakończenie podamy jeszcze kilka uwag ogólnych o skręceniu płaszczyzny polaryzacji, czyli o tzw. *aktywności optycznej*. We wstępie mówiliśmy, że jest to zjawisko dość pospolite, wykazywane zarówno przez ciała stałe, ciekłe, gazowe, jak i przez roztwory. Zatrzymamy się chwilę na aktywności optycznej kryształów. Ogólnie znanym kryształem skręcającym płaszczyznę polaryzacji jest kwarc, przy czym istnieją w przyrodzie dwie odmiany kwarcu: jedna, skręcająca w prawo — odmiana *d*, druga, w lewo — odmiana *l*. Jest to zresztą ogólna właściwość ciał optycznie czynnych, że występują one w dwóch odmianach *enancjomorficznych*, tzn. jedna odmiana różni się od drugiej tak, jak różnią się przedmiot i jego obraz w zwierciadle płaskim. *Ciało* jest *prawoskrętne*, lub *lewoskrętne*, jeśli po wstawieniu go między skrzyżowane nikole przez obrót analizatora w prawo lub w lewo uzyskuje się ponowne zupełne zaciemnienie pola widzenia. Właściwości skręcające takich kryształów jak kwarc wiążą się z asymetrią ułożenia

cząsteczek w elementarnej komórce kryształu (z brakiem płaszczyzny symetrii i środka symetrii), gdyż przy zniszczeniu sieci krystalicznej przez stopienie kryształu zanika zdolność skręcania płaszczyzny polaryzacji.

Kąt skręcenia θ płaszczyzny polaryzacji przez płytki krystaliczne przy przebiegu promieni równoległych do osi optycznej (płytka wycięta prostopadle do osi optycznej) jest proporcjonalny do grubości płytki d:

$$\theta = \alpha d,$$

gdzie α jest *skręceniem właściwym* (lub *zdolnością skręcającą*) danego ciała. Oznacza ono kąt skręcenia płaszczyzny polaryzacji wywołany przejściem promienia przez płytkę o grubości 1 mm. Skręcenie właściwe jest w dużym stopniu zależne od rodzaju ciała optycznie czynnego, jak również od długości fali (dla krótszych fal skręcenie właściwe jest większe). Zależność skręcenia właściwego od długości fali w przypadku kwarcu widzimy w tab. 26.2.

Tabela 26.2

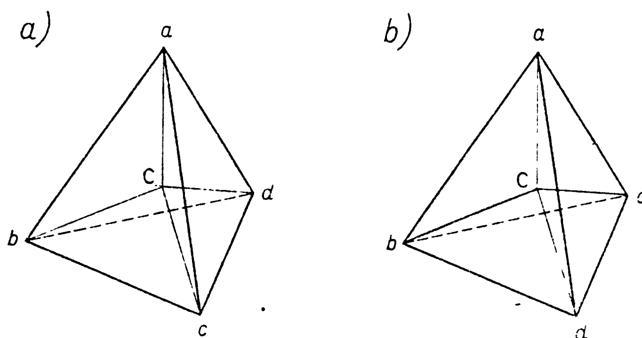
Skręcenie właściwe kwarcu w zależności od długości fali

Długość fali Å	Skręcenie właściwe stopień/mm
7608,2	12,704
6867,2	15,742
6562,819	17,314
5892,98	21,72
4861,35	32,76
4340,49	41,92
4307,76	42,63

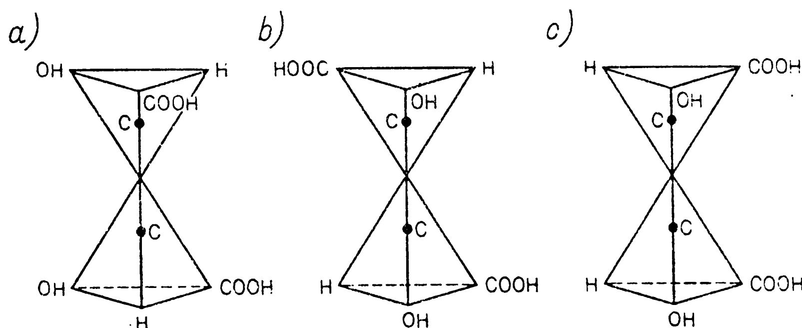
Wskutek zależności skręcenia właściwego od długości fali w świetle wychodzącym z płytki kwarcowej wyciętej prostopadle do osi, umieszczonej między nikolami i oświetlonej światłem białym, powstają efekty barwne.

Aktywność optyczną kryształów, takich jak kwarc, przypisaliśmy pewnej asymetrii ułożenia cząsteczek w elementarnej komórce kryształu. Inaczej jest w przypadku np. czystych cieczy skręcających, takich jak terpentyna czy nikotyna lub w przypadku ciał rozpuszczonych w cieczach nieskręcających, takich jak woda, benzen czy alkohol etylowy. Tym razem zdolność skręcania musimy przypisywać budowie samych cząsteczek. Tak właśnie tłumaczy się aktywność optyczną roztworów cukru. Cukier jest kryształem molekularnym; zdolność skręcania płaszczyzny polaryzacji jest właściwością jego cząsteczek. Zachowują one tę zdolność po przeprowadzeniu cukru w stan roztworu. Dzięki temu właśnie możliwe jest praktyczne wyznaczenie procentowej zawartości cukru w roztworze z pomiaru kąta skręcania płaszczyzny polaryzacji. Każda z cząsteczek znajdujących się na drodze promienia świetlnego w cieczy przyczynia się do skręcenia płaszczyzny polaryzacji, a zatem ogólne skręcenie jest proporcjonalne do stężenia roztworu i do długości drogi promienia w roztworze (wzór (26.7)).

Większość związków organicznych optycznie czynnych zawdzięcza tę właściwość tzw. asymetrycznemu atomowi węgla. Atomem asymetrycznym węgla nazywamy atom związany z czterema różnymi podstawnikami. Przedstawiając przestrzennie układ wartościowości węgla otrzymamy czworościan foremny, w którego środku znajduje się właśnie atom C (rys. 26.26). Symbole a, b, c, d oznaczają różne atomy lub grupy atomów. Przy takim układzie możliwe są dwie odmiany izomeryczne, z których jedna skręca płaszczyznę polaryzacji światła w lewo, a druga — w prawo. W prostych przypadkach można nawet ustalać, jaka struktura odpowiada danemu zwrotowi skręcenia płaszczyzny polaryzacji. Zbadano też m. in. dokładnie kwas winowy mający w cząsteczce dwa jednakowe asymetryczne atomy węgla. W tym przypadku możliwe są ugrupowania przestrzenne atomów i grup atomowych, takie jak na rys. 26.27a i b. Struktura przedstawiona na rys. 26.27a odpowiada kwasowi winowemu prawoskrętnemu, a na rys. 26.27b — lewoskrętnemu. Występujący w przyrodzie kwas gronowy o takim samym wzorze jest mieszaniną kwasów winowych lewoskrętnych i prawoskrętnych w równych ilościach.



Rys. 26.26



Rys. 26.27

Prócz kwasu winowego lewoskrętnego i prawoskrętnego występuje w przyrodzie odmiana o tym samym składzie, lecz nie wykazująca aktywności optycznej. Jest to

kwas mezowinowy. Budowę cząsteczki tego kwasu przedstawia rys. 26.27c. Obie połowy cząsteczki kwasu mezowinowego skręcają płaszczyznę drgań w przeciwnych kierunkach, stąd skręcenie wypadkowe równa się zeru.

Na zakończenie tego punktu warto wspomnieć o innych możliwościach skręcania płaszczyzny polaryzacji. Na przykład w r. 1846 Faraday stwierdził, że każde ciało izotropowe przezroczyste umieszczone w silnym polu magnetycznym skręca płaszczyznę polaryzacji światła rozchodzącego się w kierunku równoległym do linii sił pola magnetycznego. Skręcenie to znika bezpośrednio po usunięciu pola.

26.7. Polaryzacja eliptyczna i kołowa. Interferencja w świetle spolaryzowanym

W paragrafie 26.6 wyjaśniliśmy działanie półfalówki. Ustaliliśmy, że dzięki różnicy prędkości promienia zwyczajnego i nadzwyczajnego po przejściu przez cienką płytkę krystaliczną powstaje między promieniami różnica faz. W przypadku półfalówki różnica ta wynosi π dla określonego promieniowania monochromatycznego. W wyniku interferencji obu promieni powstaje znów światło liniowo spolaryzowane, lecz o przesuniętym o kąt 2γ kierunku drgań.

Odrzućmy obecnie założenie, że grubość płytki kwarcowej jest tak dobrana, że zapewnia różnicę faz promieni wychodzących równą π. Niech grubość cienkiej płytki kwarcowej wywołuje dowolną różnicę faz (między zerem i 2π) byle nierówną $\pi/2$, π lub 2π. Tym razem wskutek interferencji promienia zwyczajnego i nadzwyczajnego, wychodzących z płytki, powstaje na ogół drganie eliptyczne (zgodnie z rozważaniami w § 26.1). W warunkach statycznych koniec wektora **E** opisywałby elipsę. Ponieważ jednak fala świetlna rozchodzi się w kierunku wektora **k**, więc koniec wektora **E** opisuje linię śrubową eliptyczną. W tym przypadku mówimy o *polaryzacji eliptycznej*. Natomiast istnienie różnicy faz $\pi/2$ przy narzuceniu dodatkowego warunku, że amplitudy drgań w promieniu zwyczajnym i nadzwyczajnym są jednakowe, warunkuje wystąpienie *polaryzacji kołowej*. Z tym ostatnim przypadkiem mamy do czynienia przy użyciu tzw. ćwierćfalówki i przy zastosowaniu światła padającego liniowo spolaryzowanego o kierunku drgań **E** tworzącym kąt 45° z kierunkiem osi optycznej płytki. Ten rodzaj światła padającego na płytkę zapewnia jednakowe amplitudy drgań wektora **E** w promieniu zwyczajnym i nadzwyczajnym w płytce. Interferencja obu drgań składowych po wyjściu z płytki powoduje — w związku z ruchem fali świetlnej — opisywanie przez koniec wektora **E** linii śrubowej kołowej.

Warto jeszcze podkreślić, że polaryzacji kołowej i eliptycznej nie można wykryć stosując obrót analizatora dokoła kierunku rozchodzenia się światła jako osi. Łatwo wykazać, że w przypadku światła spolaryzowanego kołowo natężenie światła przechodzącego jest stale jednakowe (jak w świetle naturalnym), a w przypadku polaryzacji eliptycznej ulega okresowym zmianom nie osiągając jednak pełnego wygaszenia (analogicznie jak przy świetle częściowo spolaryzowanym). Do wykrycia polaryzacji kołowej potrzebna jest dodatkowa ćwierćfalówka, która wprowadza dodatkową różnicę faz obu składowych $\pm\pi/2$, a zatem sprowadza całkowitą różnicę faz do 0 lub π. W obu przypad-

kach przy złożeniu obu składowych powstaje światło spolaryzowane liniowo, które może być całkowicie wygaszone przy obrocie analizatora.

W przypadku światła spolaryzowanego eliptycznie sprawa się komplikuje, ponieważ mogą tu wchodzić w grę różne różnice faz.

26.8. Polaryzacja chromatyczna

Zjawisku interferencji w świetle spolaryzowanym mogą towarzyszyć piękne efekty barwne. Rozważymy je na przykładzie płytki krystalicznej jednoosiowej, wyciętej równolegle do osi optycznej wstawionej prostopadle w drogę promieni świetlnych. Wiemy już, że dla światła monochromatycznego o długości fali λ (por. przykład 1 w § 26.6) różnica dróg optycznych δ wychodzącej fali zwyczajnej i nadzwyczajnej wyrażona w długościach fali wynosi

$$\delta = \frac{d(n_e - n_o)}{\lambda},$$

a różnica faz Δ:

$$\Delta = 2\pi \frac{d(n_e - n_o)}{\lambda}.$$

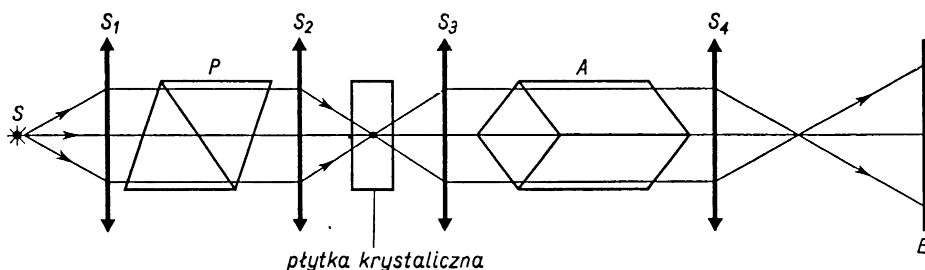
Jak widać z tych wzorów, zarówno δ, jak i Δ są — przy stałej grubości płytki d — funkcjami długości fali. Stąd wniosek, że jeśli na płytkę rzucimy światło białe, to dla poszczególnych składników tego światła otrzymamy przy przejściu przez płytkę różne przesunięcia fazowe promienia zwyczajnego i nadzwyczajnego. W związku z tym zmieniać się będą amplitudy drgań **E**, a zatem i natężenia przepuszczanych przez analizator składników. Zmiana stosunku natężeń składników światła białego powoduje wystąpienie zabarwienia. W przypadku jednakowej grubości całej płytki każdemu położeniu analizatora odpowiada określona barwa całego pola widzenia. Barwa ta ulega zmianie przy obrocie analizatora. Zmiana barwy przy danym ustawieniu analizatora świadczy o zmianie grubości płytki, np. przy wzroście kryształu. Jeśli grubość płytki nie jest wszędzie jednakowa, to otrzymujemy barwne prążki interferencyjne, przy czym prążek jednej barwy wytworzony jest przez promienie przechodzące przez części płytki o jednakowej grubości.

Barwne prążki interferencyjne można też otrzymać rzucając na płytkę krystaliczną wiązkę zbieżnych promieni. Układ doświadczalny, przedstawiony na rys. 26.28, nie wymaga bliższych objaśnień.

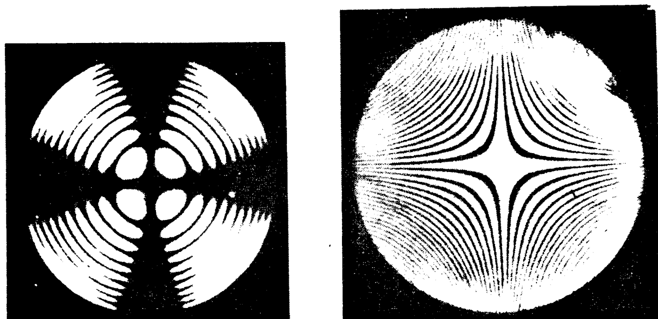
Nie wchodząc bliżej w dość skomplikowane szczegóły zjawiska ograniczymy się do podania wyników doświadczalnych dotyczących płytki wyciętej prostopadle do osi optycznej. Na ekranie powstają barwne figury interferencyjne przedstawione przykładowo na rys. 26.69. Są to pierścienie współśrodkowe, na przemian jasne i ciemne przy świetle monochromatycznym lub barwne w świetle białym, przecięte ciemnym krzyżem, którego ramiona odpowiadają wyróżnionym w doświadczeniu kierunkom drgań przepuszczanych przez analizator i polaryzator. Powstanie tych figur wiąże się

z tym, że każdemu kątowi nachylenia promieni w padającej wiązce zbieżnej odpowiada inna droga optyczna promienia zwyczajnego i nadzwyczajnego w krysztale, a więc i inna różnica faz wychodzących z płytki promieni.

Trzeba podkreślić, że efekty polaryzacji chromatycznej występują już przy bardzo nieznacznej dwójłomności optycznej. Można więc na badaniu tych efektów zarówno w świetle równoległym, jak i zbieżnym oprzeć badanie dwójłomności.



Rys. 26.28



Rys. 26.29

26.9. Dwójłomność wymuszona

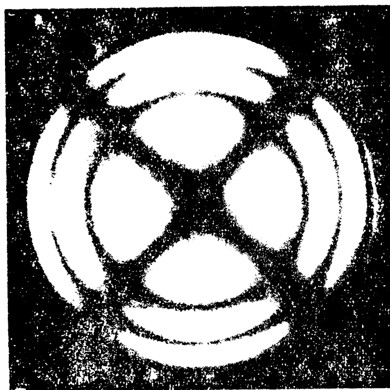
Charakteryzując zjawisko podwójnego załamania podkreśliliśmy, że występuje ono w ciałach anizotropowych optycznie, m.in. w kryształach, z wyjątkiem kryształów układu regularnego. Aby zbytnio nie ograniczyć zagadnienia, musimy jeszcze wspomnieć o anizotropii wymuszonej.

Większość ciał ma charakter anizotropowy, gdyż olbrzymia większość cząsteczek chemicznych wykazuje anizotropię budowy.

W przypadku całkowicie bezładnego, przypadkowego rozkładu przestrzennego olbrzymiej liczby takich cząsteczek, stanowiących ciało makroskopowe, anizotropia się nie ujawnia; badany obiekt zachowuje się tak, jak ciało izotropowe. Jeżeli jednak oddziaływania zewnętrzne spowodują uporządkowanie rozkładu cząsteczek, czyli wystąpienie uprzywilejowanych kierunków — ujawni się to pod postacią anizotropii optycznej.

Okazuje się na przykład, że można wywołać anizotropię optyczną i towarzyszącą jej dwójłomność przez oddziaływania mechaniczne (ściskanie, rozciąganie itd.), przez

oddziaływania termiczne (nierównomierne ogrzewanie masy ciała), przez oddziaływanie pola elektrycznego (tzw. zjawisko Kerra w cieczach), przez spowodowanie przepływu cieczy itp. Zatrzymamy się chwilę nad dwójłomnością optyczną przezroczystych ciał stałych (szkło, masy plastyczne) poddanych mechanicznym działaniom. Płytki z takich ciał zachowują się na ogół jak kryształy jednoosiowe. Taka odkształcona płytka wprowadzona między skrzyżowane nikole daje barwne prążki interferencyjne, których kształt jest funkcją rozkładu naprężeń w badanej płytce (rys. 26.30). A więc analizę rozkładu naprężeń można w ten sposób sprowadzić do badań optycznych. Metoda ta znalazła praktyczne zastosowania. Z przezroczystej masy plastycznej formuje się dokładny model jakiejś konstrukcji, np. mostu, poddaje się model przewidzianym działaniom mechanicznym i bada się efekty dwójłomności wymuszonej (elastometria optyczna).



Rys. 26.30

Pytania i zadania

W tym niewątpliwie dość trudnym rozdziale szereg pytań kontrolnych i zadań umieszczono bezpośrednio w tekście. Warto do nich wrócić raz jeszcze po przerobieniu całości rozdziału.

1. Jakich wiadomości o naturze światła dostarcza odkrycie zjawisk ugięcia, interferencji i polaryzacji?

2. Do czego sprowadza się istota polaryzacji liniowej światła? Polaryzację liniową nazywa się czasem polaryzacją płaską. Uzasadnić obie nazwy.

3. Omówić dokładnie różnicę między światłem niespolaryzowanym i światłem całkowicie liniowo spolaryzowanym. Dlaczego światło liniowo spolaryzowane i niespolaryzowane różnią się symetrią względem promienia świetlnego?

4. Czy istnieją powiązania między światłem liniowo spolaryzowanym i światłem: a) kołowo spolaryzowanym, b) eliptycznie spolaryzowanym?

5. Jakie są metody wywoływania częściowej i całkowitej polaryzacji światła?

6. Czym się różni stan polaryzacji promienia odbitego pod kątem Brewstera α_B od płytki szklanej i promienia, który przeszedł przez nią i następnie przez kilka zestawionych równolegle takich samych płytek ($\alpha = \alpha_B$). Który z promieni ma większe natężenie?

7. W jakich warunkach występuje zjawisko podwójnego załamania światła? Jakie prawa rządzą każdym z powstających promieni? Wyjaśnić szczegółowo rys. 26.16.

8. Jaki jest stan polaryzacji promienia zwyczajnego i nadzwyczajnego?

9. Wymienić kilka urządzeń polaryzacyjnych i omówić ich budowę, działanie i zastosowania.

10. Na czym polega działanie półfalówki? Dlaczego znalazła ona zastosowanie np. w sacharymetrze Laurenta? Do czego służy ten przyrząd?

11. Jakie ciała wykazują aktywność optyczną? Wyjaśnić zjawisko prawoskrętności i lewoskrętności.

12. Na czym polega polaryzacja chromatyczna?

13. Jakie zjawisko nosi nazwę dwójłomności wymuszonej? Podać przykłady i zastosowania.

14. Czy w dzisiejszych czasach badanie zjawiska polaryzacji ma tylko znaczenie historyczne, czy też ma podstawowe znaczenie w wielu dziedzinach nauki i techniki?

15. Jaki jest kąt Brewstera i odpowiadający mu kąt załamania dla płytki szklanej o $n = 1{,}5$?

Odp. 56,3°; 33,7°.

16. Znaleźć kąt Brewstera charakterystyczny dla odbicia na granicy szkło-woda ($n_s = 1{,}5$; $n_w = 1{,}33$)

Odp. 41,5°.

17. Na szklaną płytkę płasko-równoległą umieszczoną w powietrzu pada promień pod kątem Brewstera. Wykazać, że promień odbity od drugiej powierzchni tej płytki jest również całkowicie spolaryzowany.

18. Pod jakim kątem powinien padać promień świetlny na płaszczyznę rozgraniczającą wodę i szkło, by promień odbity od szkła był całkowicie liniowo spolaryzowany?

Odp. około 48°.

19. Stwierdzono, że przy padaniu światła naturalnego na pewną płytkę płasko-równoległą pod kątem Brewstera 10% promieniowania ulega odbiciu. Obliczyć, jaki procent natężenia promieniowania padającego stanowią natężenia składowych spolaryzowanych i niespolaryzowanych w: a) promieniu załamanym w płytce, b) promieniu odbitym od drugiej powierzchni płytki, c) promieniu wychodzącym z płytki.

Odp. a) 10%, 80%, b) 8%, c) 18%, 64%.

20. Jak zmniejszy się natężenie światła po przejściu przez dwa nikole zestawione tak, że kąt między ich płaszczyznami polaryzacji wynosi 60°. Każdy z nikoli z osobna zmniejsza natężenie o 10%.

Odp. 10 razy.

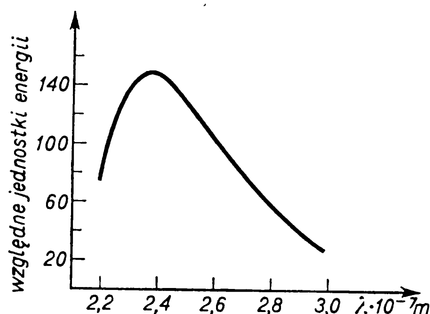
ROZDZIAŁ 27

Fotometria

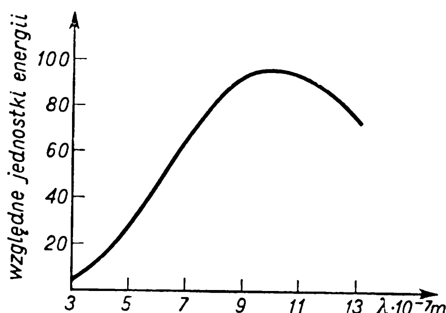
27.1. Rozróżnienie fotometrii energetycznej i wizualnej

Przedmiotem fotometrii są pomiary pewnych wielkości związanych z promieniowaniem. Jedną z podstawowych wielkości fotometrycznych jest całkowita moc promieniowania danego źródła, wyrażana w watach.

Zastanówmy się bliżej nad tym pojęciem. Czy rzeczywiście całkowita ilość energii wysyłanej w jednostce czasu przez dane źródło daje pełną charakterystykę jego właściwości świetlnych? Weźmy pod uwagę zwykłą żarówkę z drucikiem wolframowym i powierzchnię małego żelazka elektrycznego. Żarzenie się drucika żarówki odbywa się kosztem energii prądu elektrycznego. Około 20% dostarczanej energii stanowią straty



Rys. 27.1



Rys. 27.2

związane z przewodzeniem i unoszeniem ciepła. Resztę drucik wypromieniowuje w postaci fal elektromagnetycznych. Powierzchnia rozgrzanego żelazka jest również źródłem promieniowania. Można by tak dobrać oba źródła promieniowania, aby wysyłane przez nie ilości energii promienistej w czasie jednej sekundy były jednakowe. Wartość ich jednak jako źródeł światła, tzn. źródeł promieniowania działającego na zmysł wzroku, jest bardzo różna: żarówka jest źródłem światła, a żelazko nim nie jest.

Weźmy inne zestawienie źródeł światła: lampę wodorową o promieniowaniu bo-

gatym w nadfiolet i lampę wolframową o temperaturze włókna około 2700K. Spektralny rozkład energii wysyłanej przez oba źródła przedstawiają rys. 27.1 i 27.2. Widać wyraźnie walory lampy wodorowej jako źródła promieniowania niewidzialnego, bogatego w nadfiolet, i lampy wolframowej — podstawowego źródła promieniowania widzialnego.

Wymienione przykłady uzasadniają rozróżnienie fotometrii energetycznej i wizualnej, zwanych też czasem fotometrią obiektywną i subiektywną. *Fotometria energetyczna* zajmuje się całkowitą energią promienistą (widzialną i niewidzialną) i wielkościami z nią związanymi. *Fotometria wizualna* ogranicza się do badania wielkości związanych z promieniowaniem widzialnym. W związku z tym w obu dziedzinach fotometrii stosowane są różne odbiorniki energii promienistej. W fotometrii energetycznej stosowane są odbiorniki nieselektywne, tzn. reagujące na wszelkie rodzaje promieniowania. Wskazania takich odbiorników, jak np. termoogniw z poczernionymi spojeniami, są proporcjonalne do całkowitej mocy padającego promieniowania, niezależnie od reprezentowanych w tym promieniowaniu długości fal.

W fotometrii wizualnej podstawowym odbiornikiem jest nasze oko — odbiornik selektywny, reagujący tylko na zakres promieniowania widzialnego. Do odbiorników selektywnych zaliczymy też fotokomórki (§ 29.7), kliszę fotograficzną, fotopowielacze (§ 29.7), detektory półprzewodnikowe (punkt 32.14.1) itd.

W fotometrii energetycznej i wizualnej zajmujemy się podobnymi pojęciami. Mówimy o strumieniu energii promieniowania i strumieniu świetlnym, o gęstości kątowej strumienia energii promieniowania i gęstości kątowej strumienia świetlnego (światłości), o natężeniu napromieniowania i o natężeniu oświetlenia, o luminancji energetycznej i luminancji. Wielkości wymienione na pierwszym miejscu — są to podstawowe wielkości fotometrii energetycznej, wielkości wymienione na drugim miejscu — są to podstawowe wielkości fotometrii wizualnej. Definicje odpowiadających sobie wielkości są bardzo podobne, toteż będziemy stosowali w fotometrii energetycznej i wizualnej te same symbole, dodając tylko indeksy „*e*" w fotometrii energetycznej. Jednostki stosowane w obu dziedzinach fotometrii różnią się zasadniczo.

27.2. Podstawowe wielkości fotometrii energetycznej

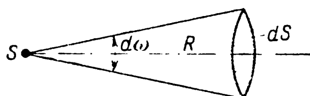
Omówimy kolejno następujące wielkości: 1) strumień energii promieniowania, 2) gęstość kątową strumienia energii promieniowania, 3) natężenie napromieniowania, 4) luminancję energetyczną, odpowiadającą danemu kierunkowi wysyłanego promieniowania.

1. Całkowity *strumień energii promieniowania* Φ_e wyraża moc danego źródła, tzn. ogólną ilość energii promienistej wysyłanej przez źródło we wszystkich kierunkach w jednostce czasu. *Jednostką strumienia energii promieniowania* jest wat (W).

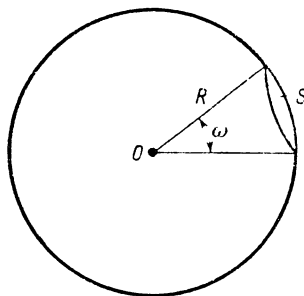
Oprócz pojęcia całkowitego strumienia energii promieniowania stosuje się w fotometrii energetycznej pojęcie strumienia energii promieniowania $d\Phi_e$, przechodzącego przez daną powierzchnię. Niech na drodze promieniowania idącego od punktowego źródła S znajduje się powierzchnia dS (rys. 27.3). Stosunek ilości energii promienistej

przechodzącej w czasie dt przez powierzchnię dS do czasu dt jest miarą strumienia energii promieniowania przechodzącego przez powierzchnię dS.

Strumień energii promieniowania można też uzależnić od kąta bryłowego, w jakim rozchodzi się promieniowanie. Na przykład do zaznaczonej na rys. 27.3 powierzchni dS dochodzą ze źródła S tylko promienie zawarte w stożku opartym na konturze dS. Takiemu stożkowi odpowiada kąt bryłowy $d\omega$. Przypominamy określenie kąta bryłowego i jego jednostkę. Wyobraźmy sobie kulę o promieniu R opisaną z punktu O (rys. 27.4). Na powierzchni tej kuli „opieramy" stożek o wierzchołku w punkcie O. Stożek ten „opiera się " na polu S wyciętym z tej kuli, a równocześnie wyodrębnia część przestrzeni, której przypisujemy kąt bryłowy ω. Miarą tego kąta jest stosunek pola S wyciętego przez stożek na kuli do kwadratu jej promienia. Jednostką kąta bryłowego jest *steradian* (skrót sr). Jednostkowy kąt bryłowy opiera się na polu 1 m^2 wyciętym z kuli o promieniu równym 1 metrowi. Pełny kąt bryłowy wynosi 4π steradianów.



Rys. 27.3



Rys. 27.4

Po tej krótkiej dygresji wróćmy do omawianego pojęcia strumienia energii promieniowania. Zgodnie z rys. 27.3 strumień rozchodzi się w stożku odpowiadającym kątowi bryłowemu $d\omega$, czyli krótko mówiąc — strumień $d\Phi_e$ przypada na kąt bryłowy $d\omega$. Jeśli ośrodek nie pochłania rozchodzącej się w nim energii promieniowania, to przez dowolne przecięcie omawianego stożka powierzchniami kulistymi o środku w punkcie S przechodzi taki sam strumień energii promienistej.

Energia promienista rozchodzi się od badanego punktowego źródła S we wszystkich kierunkach. Jeśli strumienie energii nie zależą od kierunku rozchodzenia się energii, mówimy o *równomiernym strumieniu energii promieniowania* we wszystkich kierunkach.

2. Następna wielkość fotometryczna, tzw. gęstość kątowa strumienia energii promieniowania wiąże się z poprzednią. Przez *gęstość kątową strumienia energii promieniowania* I_e w danym kierunku rozumiemy stosunek strumienia energii do wielkości kąta bryłowego, jakiemu on odpowiada:

$$I_e = \frac{d\Phi_e}{d\omega}. \tag{27.1}$$

W przypadku równomiernego strumienia gęstość kątową energii promieniowania otrzymujemy, dzieląc wartość całkowitego strumienia przez pełny kąt bryłowy:

$$I_e = \frac{\Phi_e}{4\pi}.$$

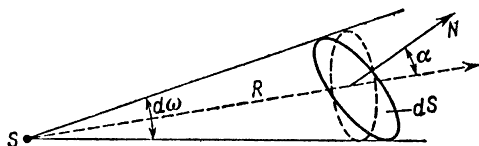
Jednostką gęstości kątowej energii promieniowania jest wat na steradian (W/sr):

Rozmaite urządzenia, wykorzystujące przede wszystkim ogniskowanie promieni przy załamaniu i odbiciu, mają na celu właśnie zwiększanie gęstości kątowej strumienia energii promieniowania w pewnych kierunkach, kosztem zmniejszania jej w kierunkach pozostałych.

3. W punkcie 1 omówiliśmy strumień energii promienistej $d\Phi_e$ przechodzący przez powierzchnię dS. Stosunek

$$E_e = \frac{d\Phi_e}{dS} \tag{27.2}$$

jest miarą następnej wielkości fotometrycznej, którą nazywamy natężeniem napromieniowania. *Natężenie napromieniowania* E_e wyraża liczbowo strumień energii przypadający na jednostkę powierzchni. *Jednostką natężenia napromieniowania* jest W/m^2.



Rys. 27.5

Wzór określający natężenie napromieniowania możemy przekształcić wprowadzając do niego odległość od punktowego źródła energii promienistej. Przede wszystkim zamiast dowolnie ustawionego elementu powierzchni dS bierzemy jego rzut na płaszczyznę prostopadłą do kierunku rozchodzenia się promieniowania. Jeśli normalną do dS tworzy kąt α z kierunkiem rozchodzenia się promieniowania (rys. 27.5), to szukany rzut wynosi $dS \cos\alpha$. Na tej powierzchni opiera się kąt bryłowy $d\omega$, który można wyrazić jako

$$d\omega = \frac{dS\cos\alpha}{R^2}.$$

Ponieważ równocześnie spełnione jest równanie

$$I_e d\omega = d\Phi_e,$$

więc natężenie napromieniowania E_e równa się

$$E_e = \frac{I_e \cos\alpha}{R^2}. \tag{27.3}$$

W ten sposób doszliśmy do jednego z najważniejszych wzorów fotometrycznych, uzależniających natężenie napromieniowania E_e na danej powierzchni od jej odległości R od

źródła punktowego, od gęstości kątowej I_e strumienia energii i od kąta α zawartego między kierunkiem normalnej do badanej powierzchni i kierunkiem rozchodzenia się promieniowania, wyznaczonym przez oś kąta bryłowego $d\omega$.

4. Pozostała jeszcze do omówienia wielkość czwarta, tzw. luminancja energetyczna, związana z rozciągłym (nie punktowym) źródłem, odpowiadająca danemu kierunkowi rozchodzenia się promieniowania.

Obecnie element dS wyodrębniamy z powierzchni źródła. Normalna do tego elementu tworzy kąt α z rozważanym kierunkiem rozchodzenia się promieniowania. Można uważać, że promieniowanie rozchodzi się prostopadle do elementu powierzchni $dS\cos\alpha$. Wtedy *luminancją energetyczną* źródła w kierunku α nazywamy stosunek

$$L_e = \frac{d\Phi_e}{d\omega dS\cos\alpha}, \tag{27.4}$$

czyli

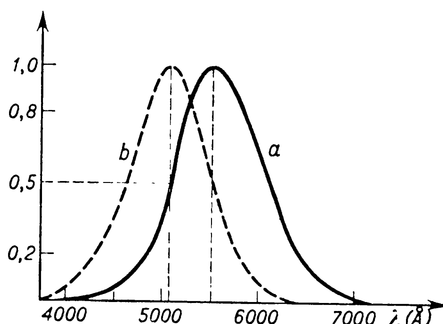
$$L_e = \frac{dI_e}{dS\cos\alpha}. \tag{27.5}$$

Jednostką luminancji energetycznej jest W/sr $\cdot$ m^2.

Tylko ciała doskonale czarne mają luminancję energetyczną niezależną od kierunku. Niektóre źródła, m.in. Słońce, można w przybliżeniu traktować jako ciała o luminancji energetycznej niezależnej od kierunku.

27.3. Krzywa wrażliwości oka ludzkiego

Przechodzimy obecnie do fotometrii wizualnej. Odbiornikiem energii światła widzialnego jest przede wszystkim oko ludzkie, dlatego też nawiążemy tu do jego właściwości.



Rys. 27.6

Oko ludzkie reaguje na długości fal w przedziale mniej więcej 4000–8000 Å (czyli $4\cdot10^{-7}$–$8\cdot10^{-7}$ m). Wrażliwość oka nie jest jednak w całym tym zakresie jednakowa. Ponieważ ludzie o normalnym wzroku podobnie reagują na fale o różnych długościach, można więc nakreślić krzywą przedstawiającą w pewien umowny sposób wrażliwość przeciętnego oka. Na osi odciętych (rys. 27.6) odkładamy długości fali, na osi rzędnych wartości zawarte między 1 i 0. Oko ludzkie adaptowane „na jasność” (krzywa *a*) wy-

kazuje największą wrażliwość na promieniowanie o długości 5550 Å. Tej długości fali w sposób umowny przypisujemy rzędną jednostkową na wykresie. Wszystkim pozostałym długościom fal, ze względu na mniejszą wrażliwość oka, przypisujemy rzędne ułamkowe. Na przykład długości fali 5100 Å przypisujemy rzędną $\frac{1}{2}$. Odwrotność tej rzędnej wskazuje, ile razy więcej energii promienistej odpowiadającej fali o długości 5100 Å należy doprowadzić do oka, aby wywołać takie samo wrażenie jasności, jak przy użyciu fali o długości 5550 Å. Dla promieni ciemnoczerwonych o długości 7100 Å rzędna wynosi $\frac{1}{500}$, czyli do wywołania wrażenia tej samej jasności potrzebna jest 500 razy większa ilość energii. Zagadnienie zresztą nie jest zbyt proste, gdyż trudno jest oceniać równość wrażeń jasności, gdy wchodzą w grę światła o różnych barwach.

Przebieg krzywej wrażliwości oka ludzkiego adaptowanego „na ciemność" (krzywa *b*) jest nieco inny. Maksimum krzywej przesunięte jest w stronę fal krótkich i odpowiada długości 5100 Å. Przy słabym oświetleniu zmniejsza się wyraźnie wrażliwość na promieniowanie długofalowe. Wiemy o tym z codziennych doświadczeń: o zmroku przedmioty barwy czerwonej wydają się prawie czarne.

Na podstawie tych kilku uwag można wnioskować, że obiektywna ocena źródeł promieniowania, oparta na wrażeniach wzrokowych, jest niezmiernie trudna i wymaga uwzględniania wielu czynników, takich jak indywidualne właściwości oka obserwatora, adaptacja oka „na jasność" lub „na ciemność" oraz skład promieniowania dochodzącego do oka.

27.4. Podstawowe wielkości i jednostki fotometrii wizualnej

Zgodnie z wprowadzoną w § 27.1 umową nie powtarzamy już definicji takich wielkości jak strumień świetlny, światłość (czyli gęstość kątowa strumienia świetlnego), natężenie oświetlenia, luminancja. Definicje te byłyby powtórzeniem definicji z dziedziny fotometrii energetycznej, z tym zastrzeżeniem, że odnosiłyby się tylko do energii promieniowania widzialnego.

Wielkości wyżej wymienione wyrażane są w specjalnych jednostkach różnych od jednostek stosowanych w fotometrii energetycznej. Podstawową jednostką fotometrii wizualnej jest *jednostka światłości*, zwana *kandelą* (skrót cd), będąca równocześnie jedną z podstawowych jednostek układu SI.

Światłość równą *jednej kandeli* ma w kierunku prostopadłym pole $\frac{1}{6\cdot 10^5}$ m^2 powierzchni ciała doskonale czarnego, promieniującego w temperaturze krzepnięcia platyny pod ciśnieniem 101 325 N/m^2 (czyli 1 atmosfery fizycznej).

Opierając się na równaniu definiującym światłość (czyli gęstość kątową strumienia świetlnego lub, inaczej mówiąc, strumienia energii promieniowania widzialnego):

$$I = \frac{d\Phi}{d\omega}, \tag{27.6}$$

możemy ustalić wzór na strumień świetlny $d\Phi$:

$$d\Phi = I d\omega \tag{27.7}$$

i następnie przejść do definicji *jednostki strumienia świetlnego* zwanej *lumenem* (skrót lm):

$$1 \text{ lm} = 1 \text{ cd} \cdot 1 \text{ sr}.$$

Należy to rozumieć w ten sposób: gdy punktowe źródło światła ma we wszystkich kierunkach światłość równą 1 kandeli, to na każdy steradian przypada strumień świetlny równy jednemu lumenowi. W tych warunkach całkowity strumień świetlny wysyłany przez źródło wynosi 4π lumenów, czyli 12,566 lm.

Niejednakowa wrażliwość oka ludzkiego na różne barwy światła decyduje o tym, że dla różnych źródeł monochromatycznych występuje różny stosunek odpowiadających sobie lumenów i watów. Dla światła monochromatycznego o długości 5550 Å, odpowiadającego największej wrażliwości oka, mamy najkorzystniejszy stosunek, gdyż na 1 W przypada 620 lm. Żarówka 25 W daje zaledwie około 300 lm, tzn. na 1 W przypada tylko około 12 lm. Wiąże się to z faktem, że w promieniowaniu żarówki występują w dużym procencie fale długie, na które oko słabo reaguje.

Jednostka natężenia oświetlenia wynika z równania

$$E = \frac{d\Phi}{dS}. \tag{27.8}$$

Jednostką natężenia oświetlenia jest lumen na metr kwadratowy (lm/m²). Jednostce tej nadano nazwę luks (skrót lx):

$$1 \text{ lx} = 1 \frac{\text{lm}}{\text{m}^2}.$$

Luks jest natężeniem oświetlenia wywołanego przez strumień świetlny 1 lumena, przypadający na powierzchnię 1 metra kwadratowego.

Natężenie oświetlenia równe 1 luksowi panuje na powierzchni kuli o promieniu $R = 1$ m, w środku której znajduje się punktowe źródło o światłości jednakowej we wszystkich kierunkach i równej 1 cd.

W fotometrii energetycznej uzasadniliśmy wzór (27.3). W fotometrii wizualnej stosujemy wzór analogiczny:

$$E = \frac{I \cos \alpha}{R^2}. \tag{27.9}$$

(Znaczenie symboli omówiono przy wzorze 27.3.) Ze wzoru (27.8) można odczytać, że przy *niezmiennej światłości I punktowego źródła i niezmiennym kącie α natężenie oświetlenia E maleje jak kwadrat odległości od źródła*:

$$\frac{E_1}{E_2} = \frac{R_2^2}{R_1^2}. \tag{27.10}$$

Wzór ten, aczkolwiek wyprowadzony dla źródeł punktowych, stosowany jest często w fotometrii wizualnej i do źródeł rozciągłych. Przybliżenie jest tym lepsze, im rozmiary źródła są mniejsze w porównaniu z odległością R występującą we wzorze.

Jak widać ze wzoru (27.10), natężenie oświetlenia E można zmieniać przesuwając powierzchnię oświetlaną w stosunku do źródła światła.

Odpowiedni dobór natężenia oświetlenia i zapewnienie jego równomierności jest jednym z zasadniczych czynników decydujących o wydajności pracy ludzkiej. Przy pracach niezbyt precyzyjnych wystarcza natężenie oświetlenia około 30 lx. Przy pracach bardziej precyzyjnych potrzebne są natężenia oświetlenia dochodzące do 200–300 lx. W letnie słoneczne popołudnie oświetlenie E dochodzi do 100 000 luksów.

Jednostką luminancji w fotometrii wizualnej jest nit (skrót nt). Mówimy, że luminancja źródła wynosi 1 nit, gdy powierzchnia równa 1 m^2 ma w kierunku normalnym światłość jednej kandeli. Dla orientacji podajemy, że luminancja włókna wolframowego przeciętnej żarówki wynosi około $3 \cdot 10^6$ nitów, a Słońca około $3 \cdot 10^9$ nitów.

$$1 \text{ nit} = 1 \text{ nt} = 1 \frac{\text{cd}}{\text{m}^2}.$$

27.5. Fotometry

Fotometry są to urządzenia służące do porównywania światłości różnych źródeł światła.

Stosując oko jako *bezpośredni* odbiornik światła dochodzącego z różnych źródeł nie potrafimy określić ani bezwzględnych wartości światłości, ani nawet stosunku światłości porównywanych źródeł. Można jednak wykonywać odpowiednie pomiary wykorzystując zdolność oka do wykrywania drobnych nawet różnic w natężeniu oświetlenia sąsiadujących ze sobą oświetlonych powierzchni. W praktyce fotometrycznej stosuje się zazwyczaj następujący sposób postępowania. Dwa pola stykające się ze sobą wzdłuż pewnej linii granicznej oświetlamy z dwóch źródeł (jedno z nich może być źródłem wzorcowym o znanej światłości I_1, drugie — źródłem badanym o światłości I_2). Natężenia oświetleń obu pól wyrównujemy bądź przez oddalanie silniejszego źródła od badanej powierzchni, bądź przez sztuczne osłabianie silniejszego strumienia świetlnego w znanym stosunku. Można to osiągnąć przez stosowanie odpowiednich przesłon, np. przesłon absorbujących w postaci płytek ze szkła matowego lub kwarcu, albo też w postaci tzw. filtrów szarych pokrytych srebrem lub platyną, którym nadaje się kształt klinów lub osłabiaczy schodkowych. Osłabienie wiązek uzyskuje się najczęściej przez wykorzystanie zjawiska polaryzacji (§ 26.5) stosując odpowiednie skręcanie nikoli.

W przypadku wyrównywania natężeń oświetlenia przez odpowiedni dobór odległości obu źródeł opieramy się na wzorze (27.9)

$$E_1 = \frac{I_1 \cos \alpha_1}{R_1^2}, \qquad E_2 = \frac{I_2 \cos \alpha_2}{R_2^2}.$$

Jeżeli przez odpowiednie ustawienie obu źródeł światła wyrównamy E_1 i E_2 przy jednakowych kątach α_1 i α_2 (kąty między normalną do oświetlanej powierzchni i kierunkiem promieni padających ze źródeł I i II), to

$$\frac{I_1}{I_2} = \frac{R_1^2}{R_2^2}.$$

Znając światłość źródła wzorcowego I_1 można wyznaczyć I_2.

Przepuszczając światło przez klin absorpcyjny można zmniejszać strumień świetlny w żądanym stosunku, ponieważ klin pochłania światło proporcjonalnie do swej grubości. Strumień świetlny Φ po przejściu przez warstwę absorbującą o grubości dx ulega osłabieniu o

$$-d\Phi = k\Phi dx.$$

Współczynnik proporcjonalności k nazywamy *stałą absorpcji*. Osłabienie strumienia świetlnego jest proporcjonalne do strumienia padającego i do grubości warstwy pochłaniającej. Wartość strumienia świetlnego po przejściu przez warstwę o grubości d otrzymujemy całkując równanie

$$-\frac{d\Phi}{\Phi} = k dx,$$

$$\ln\Phi = -kx + \text{const}.$$

Uwzględniając warunki graniczne (gdy $x = 0$, $\Phi = \Phi_0$) otrzymujemy

$$c = \ln\Phi_0,$$

czyli

$$\Phi = \Phi_0 e^{-kx}.$$

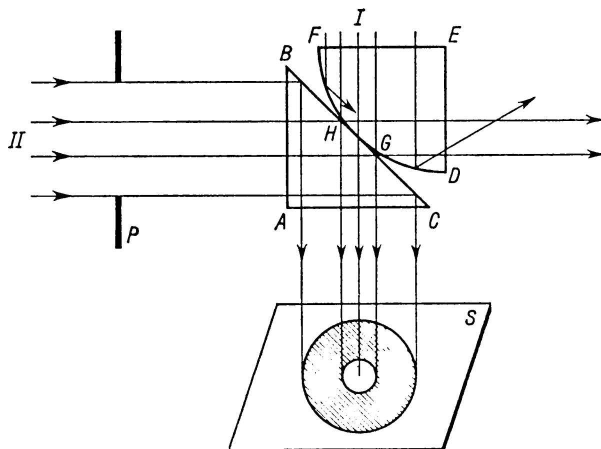
Stała absorpcji k jest na ogół zależna od długości fali, można jednak konstruować kliny o stałej absorpcji w przybliżeniu niezmiennej w całym widzialnym obszarze promieniowania.

Dla celów fotometrycznych można stosować zamiast klinów z zaczernionego szkła klisze fotograficzne, naświetlone w ten sposób, że czas naświetlania wzrasta w sposób ciągły przy przechodzeniu od jednej krawędzi kliszy do drugiej. Klisza taka po wywołaniu zawiera w emulsji koloidalny osad srebra w ilości stopniowo rosnącej przy przejściu od miejsc mniej naświetlonych do bardziej naświetlonych.

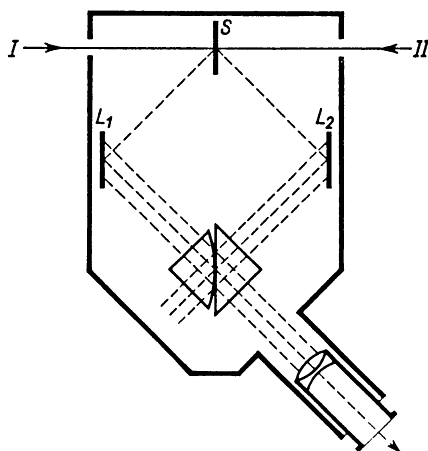
Omówimy nieco dokładniej jedną z metod fotometrycznych wyznaczania światłości dowolnego źródła, gdy rozporządzamy źródłem o światłości znanej. Posługujemy się tzw. *kostką fotometryczną Brodhuna-Lummera*, umożliwiającą obserwację dwóch przylegających do siebie pól oświetlonych światłem z dwóch różnych źródeł. Kostka składa się z dwóch prostokątnych pryzmatów całkowicie odbijających, zestawionych tak jak na rys. 27.7. Pryzmat *DEF* ma podstawę kulistą *DF*, zeszlifowaną płasko w części *GH*. Ta część jest równocześnie powierzchnią styku obu pryzmatów. Płasko zeszlifowana część *GH* ma kształt koła. Ponieważ oba pryzmaty są silnie do siebie dociśnięte, promienie świetlne idące od źródła *I* przechodzą bez przeszkód w obrębie koła styku i dają na ekranie *S* jasny krążek. Boczne promienie ulegają całkowitemu odbiciu. Z promieni idących od źródła *II* te, które trafią na koło styku, przechodzą również bez przeszkód, ale właśnie one nie dochodzą do ekranu *S*. Pozostałe promienie idące ze źródła *II* dzięki całkowitemu wewnętrznemu odbiciu skierowane są na ekran *S* i dają dokoła jasnego krążka obszar oświetlony o kształcie zależnym od kształtu użytej przesłony *P*. W przypadku przesłony kołowej otrzymamy na ekranie *S* krążek i otaczający go pierścień. Zniknięcie linii rozdzielającej krążek i pierścień, obserwowane

za pomocą lupy, odpowiada wyrównaniu natężeń oświetlenia wywołanych przez porównywane źródła. Fotometr Brodhuna–Lummera przedstawiony jest na rys. 27.8.

Zaznaczony bieg promieni wyjaśnia całkowicie zasadę działania przyrządu. S oznacza przesłonę rozpraszającą równomiernie światło dochodzące od obu źródeł. L_1 i L_2 są to płaskie zwierciadła służące do zmiany kierunku biegu promieni.



Rys. 27.7



Rys. 27.8

Zmieniając położenie jednego ze źródeł lub osłabiając w znanym stopniu strumień świetlny idący od silniejszego źródła doprowadzamy do równości natężenia oświetleń i obliczamy światłość jednego źródła znając światłość drugiego.

Pytania i zadania

1. Na czym polega różnica między fotometrią energetyczną i fotometrią wizualną?
2. Jakie są podstawowe wielkości fotometryczne? Podać ich definicje.
3. Wymienić poznane w kursie fizyki strumienie różnych wielkości.
4. Która z jednostek fotometrii wizualnej jest jedną z podstawowych jednostek układu SI? Jaki jest związek tej jednostki z jednostką strumienia świetlnego?
5. Dlaczego zależność liczbowa między lumenami i watami nie jest stała?
6. Scharakteryzować wrażliwość spektralną oka ludzkiego.
7. Omówić różne metody zmniejszania natężenia oświetlenia danej powierzchni.
8. Omówić zasadę budowy i działania fotometru Brodhuna–Lummera.
9. Obliczyć strumień świetlny wychodzący ze źródła o światłości $I = 100$ cd i padający na powierzchnię $S = 10\ \mathrm{cm}^2$ ustawioną prostopadle do kierunku padającego światła w odległości $d = 1$ m od źródła.

   Odp. 0,1 lm.
10. Obliczyć natężenie oświetlenia płaszczyzny odległej o $R = 5$ m od źródła o światłości $I = 50$ cd, jeżeli promienie padają na nią pod kątem 30°.

    Odp. 1,71 lx.
11. Nad środkiem kwadratowego stołu na wysokości $h = 2$ m wisi lampa o światłości 200 cd. Obliczyć natężenie oświetlenia w środku stołu i na jego rogach, jeżeli długość krawędzi stołu równa się 2 m.

    Odp. 50 lx; 27,7 lx.

ROZDZIAŁ 28

Rozszczepienie światła

28.1. Promieniowanie widzialne, podczerwone i nadfioletowe

W rozdziale tym, poświęconym zjawisku rozszczepienia światła, przyrządom spektralnym i analizie widmowej, będziemy się niejednokrotnie odwoływali do wiadomości zdobytych już poprzednio.

Tak więc w § 23.7 mówiliśmy o widmie fal elektromagnetycznych. Z widma tego wydzieliliśmy wąski wycinek obejmujący długości fal od $4 \cdot 10^{-7}$ m do $8 \cdot 10^{-7}$ m, jako widmo promieniowania widzialnego. Długofalowa granica tego widma to granica czerwona, krótkofalowa — to granica fioletowa. Między nimi znajdują się wszystkie widzialne barwy widmowe z sześcioma barwami zasadniczymi i około 160 różnymi odcieniami barwnymi.

Widmo promieniowania widzialnego graniczy od strony fal długich z podczerwienią sięgającą do około 10^{-4} m, a od strony fal krótkich — z nadfioletem (zwanym też ultrafioletem) sięgającym do około 10^{-8} m.

Część widzialna promieniowania wraz z sąsiadującymi z nią podczerwienią i nadfioletem występuje w widmie promieniowania cieplnego (termicznego) każdego ciała stałego ogrzanego do dostatecznie wysokiej temperatury. Można się o tym przekonać rzucając wiązkę promieniowania termicznego na ścianę pryzmatu.

Jak wiemy z § 24.11, załamaniu światła w pryzmacie towarzyszy zjawisko rozszczepienia, wynikające z zależności prędkości światła w danym ośrodku, a więc i zależności współczynnika załamania danego ośrodka od długości fali. Rzucając na ekran widmo promieniowania termicznego silnie ogrzanego ciała otrzymujemy ciągłą barwną wstęgę (tzw. widmo ciągłe), której część czerwona jest najmniej odchylona, a część fioletowa jest najbardziej odchylona od pierwotnego biegu wiązki.

Wykrycie promieniowania podczerwonego i nadfioletu wymaga użycia innych detektorów promieniowania: oko ludzkie nie spełnia tego zadania. Istnienie promieniowania podczerwonego można wykryć wykorzystując jego działanie cieplne. Można np. na drodze promieni podczerwonych ustawić pokryte sadzą spojenie termoogniwa. Wskutek ogrzania tego spojenia włączony do obwodu przyrząd pomiarowy wykazuje powstanie siły termoelektrycznej lub przepływ prądu termoelektrycznego. Można też

do tego celu stosować znane nam już bolometry lub klisze fotograficzne, uczulone specjalnie na podczerwień.

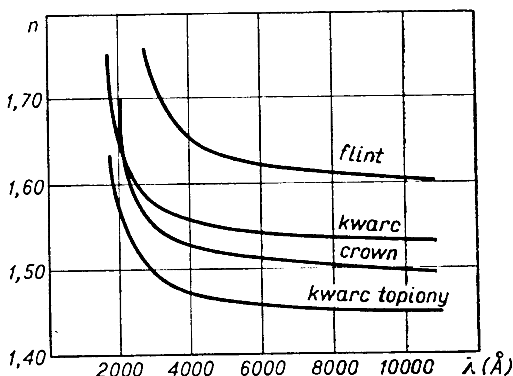
Istnienie nadfioletu można wykazać wykorzystując jego działanie chemiczne (zaczernianie kliszy fotograficznej), wzbudzanie fluorescencji i fosforescencji odpowiednio dobranych ekranów, wzbudzanie zjawiska fotoelektrycznego itp.

Badanie podczerwieni i nadfioletu w dziedzinie długości fal bardziej odległych od granic widma widzialnego nastręcza pewne trudności eksperymentalne, gdyż te rodzaje promieniowania są już silnie pochłaniane przez szkło. Nie można zatem do badań w tych dziedzinach stosować tzw. optyki szklanej, tzn. szklanych pryzmatów i soczewek. W dziedzinie nadfioletu szkło zastępuje się kwarcem lub fluorytem, a poniżej długości fal $1{,}2 \cdot 10^{-7}$ m stosuje się wyłącznie dyfrakcyjne spektrografy próżniowe. W dziedzinie podczerwieni odpowiednim nie pochłaniającym materiałem jest np. sól kamienna (NaCl), sylwin (KCl), jak również fluorek litu, bromek potasu, szafir, german.

Bogate w promieniowanie nadfioletowe jest światło pochodzące z dodatniej elektrody węglowego łuku elektrycznego, światło łuku rtęciowego, dające widmo liniowe z najsilniejszą linią o długości $\lambda = 2{,}537 \cdot 10^{-7}$ m, oraz światło lampy wodorowej.

28.2. Dyspersja normalna

W paragrafach 24.11 i 28.1 mówiliśmy już o tym, że przyczyna rozszczepienia (czyli inaczej dyspersji) światła przy przejściu przez pryzmat tkwi w zależności prędkości światła w danym ośrodku (współczynnika załamania) od długości fali. Paragraf ten poświęcimy dokładniejszemu omówieniu dyspersji.



Rys. 28.1

W przypadku wszystkich ciał przezroczystych, bezbarwnych stwierdzamy stopniowy wzrost współczynnika załamania n, gdy maleje długość fali, tzn. gdy od czerwieni przechodzimy do fioletu. W tych warunkach mówimy o tzw. *dyspersji normalnej*. Krzywe zależności współczynnika załamania od długości fali, zwane *krzywymi dyspersji*, mają w przypadku różnych ciał nieco różne przebiegi, jak to widać z rys. 28.1. Poszczególne krzywe dotyczą dwóch odmian szkła: flintu i kronu (crownu) oraz kwarcu

krystalicznego i kwarcu topionego, a więc materiałów znajdujących szerokie zastosowanie w aparaturach optycznych.

Zjawisko dyspersji normalnej można charakteryzować za pomocą kilku wielkości, jak np. dyspersji D danego materiału, dyspersji kątowej i liniowej pryzmatu, rozszczepienia całkowitego.

Przez *dyspersję D danego materiału* rozumie się wielkość

$$D = -\frac{dn}{d\lambda},$$

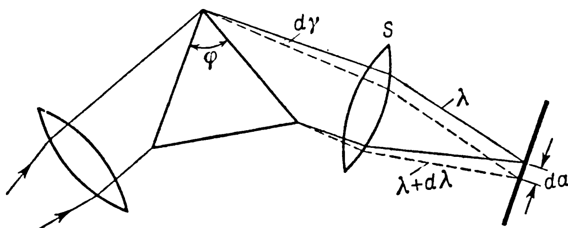
gdzie dn wyraża zmianę współczynnika załamania przy zmianie długości fali z λ na $(\lambda+d\lambda)$. Znak minus we wzorze wiąże się z tym, że ze wzrostem długości fali współczynnik załamania maleje.

Dyspersja danego materiału równa się tangensowi kąta nachylenia stycznej do krzywej dyspersji w punkcie odpowiadającym długości fali λ.

Jako miarę *dyspersji kątowej pryzmatu* przyjmujemy stosunek

$$\frac{d\gamma}{d\lambda},$$

czyli stosunek zmiany rozwartości kąta γ, pod jakim rozchodzą się po wyjściu z pryzmatu dwa promienie o mało różniących się długościach fal λ i $(\lambda+d\lambda)$ do $d\lambda$ (rys. 28.2). Można wykazać, że dyspersja kątowa pryzmatu zależy od współczynnika załamania, kąta łamiącego i od dyspersji materiału pryzmatu.



Rys. 28.2

Dyspersję liniową pryzmatu można zdefiniować na podstawie tego samego rysunku. Dyspersja liniowa wyraża się stosunkiem $da/d\lambda$, tzn. stosunkiem odległości da, w jakiej przypadają w płaszczyźnie detekcji dwie linie widmowe o długościach fal λ i $(\lambda+d\lambda)$, do $d\lambda$. Dyspersja liniowa zależy od dyspersji kątowej, jak również od ogniskowej soczewki, służącej do wytworzenia obrazu widma i od kąta nachylenia płaszczyzny detekcji względem osi wspomnianej soczewki.

Jako miarę *rozszczepienia całkowitego* przyjmuje się różnicę współczynników załamania dla skrajnych promieni fioletowych i czerwonych. Ponieważ współczynnik załamania jest funkcją długości fali, więc należy dokładnie określić, jakie długości fal uważa się za skrajny fiolet i skrajną czerwień. Zwykle wykorzystywane są do tego celu linie absorpcyjne, występujące w widmie słonecznym, oznaczone przez Fraunhofera literami B i H. Linia B leży w skrajnej czerwieni (linia A jest już na granicy widzialności), linia H — w skrajnym fiolecie (punkt 28.5.2). Przykładowo wymienimy war-

tości rozszczepienia całkowitego dla dwóch gatunków: dla szkła kron (crown) $n_H - n_B = 0{,}02073$, dla szkła flint $n_H - n_B = 0{,}04321$*.

Widma uzyskane przy użyciu pryzmatów o tym samym kącie łamiącym, lecz o różnych rozszczepieniach całkowitych, mają różne długości. Na przykład dla $\varphi = 1°$ na ekranie umieszczonym w odległości 1 m od pryzmatu długość widma wytworzonego przez pryzmat z kronu wynosi około 0,35 mm, a przez pryzmat z flintu — około 0,73 mm.

Zmieniając kąty łamiące pryzmatów można je tak dobrać, aby długości widm były jednakowe. Rozmieszczenie promieni o pośrednich barwach w obu widmach będzie jednak różne w związku z różnym charakterem zależności współczynnika załamania od długości fali w rozmaitych ciałach, czyli z różnym przebiegiem krzywych dyspersji.

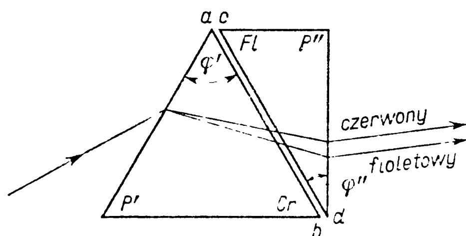
28.3. Pryzmaty achromatyczne i nie odchylające

Zajmiemy się obecnie zagadnieniem, jak zestawiać układy pryzmatów, aby:

1) wychodzące wiązki były odchylone, lecz nie rozszczepione na barwy (*układy achromatyczne*),

2) wychodzące wiązki były nie odchylone, lecz rozszczepione (*układy à vision directe*).

Wyprowadźmy *warunek achromatyczności układu dwóch pryzmatów P′ i P″* (rys. 28.3) dla skrajnych barw widma.



Rys. 28.3

Krawędzie łamiące obu pryzmatów są równoległe, ściany boczne *ab* i *cd* przylegają do siebie lub pozostają w układzie równoległym w niewielkiej od siebie odległości. Podstawy obu pryzmatów zwrócone są w przeciwnych kierunkach, tzn. oba pryzmaty odchylają promienie w przeciwne strony. Pryzmat P' (np. z kronu) ma kąt łamiący φ' i współczynnik załamania n', będący oczywiście funkcją długości fali. Symbole φ'' i n'' określają odpowiednio kąt łamiący i współczynnik załamania pryzmatu P'' (np. z flintu).

Załóżmy dla uproszczenia, że oba pryzmaty mają małe kąty łamiące. Wtedy, stosując uproszczony wzór (24.8), piszemy dla skrajnych linii Fraunhofera

$$\delta'_B = (n'_B - 1)\varphi', \qquad \delta'_H = (n'_H - 1)\varphi'.$$

* Do oznaczania dyspersji stosuje się najczęściej w optyce instrumentalnej linie Fraunhofera: czerwoną *C* i niebieską *F*.

Różnica obu odchyleń w pierwszym pryzmacie wynosi

$$\delta'_H - \delta'_B = (n'_H - n'_B)\varphi'.$$

Analogicznie, dla pryzmatu P'' różnica odchyleń w kierunku przeciwnym wynosi

$$\delta''_H - \delta''_B = (n''_H - n''_B)\varphi''.$$

Skrajne promienie będą jednakowo odchylone, gdy

$$\delta'_H - \delta'_B = \delta''_H - \delta''_B,$$

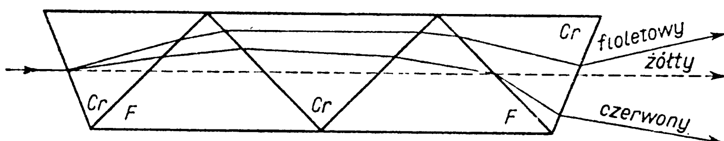
czyli gdy

$$\frac{\varphi'}{\varphi''} = \frac{n''_H - n''_B}{n'_H - n'_B}.$$

Innymi słowy, kąty łamiące obu pryzmatów muszą być odwrotnie proporcjonalne do całkowitego rozszczepienia. Wtedy przy takim ustawieniu obu pryzmatów, o jakim mówiliśmy na początku tego paragrafu, promień wychodzący jest odchylony od pierwotnego biegu, lecz barwy fioletowa i czerwona są nie rozszczepione.

Stosując dwa pryzmaty można uzyskać achromatyzację tylko dla jednej pary barw. Wiąże się to z tym, że rozszczepienie całkowite i częściowe nie są do siebie proporcjonalne w różnych pryzmatach. Chcąc otrzymać achromatyzację dla trzech barw należy już dołączyć trzeci pryzmat.

Dla ilustracji podamy kilka wartości liczbowych. Stosując pryzmat z kronu o kącie łamiącym 20° i pryzmat z flintu o kącie łamiącym 8°15′ (wartości współczynników załamania jako funkcje długości fali podane są na wykresie 28.1), można uzyskać widmo, w którym maksymalna różnica odchyleń wynosi nieco więcej niż 4′.



Rys. 28.4

W *pryzmatach nie odchylających*, czyli *à vision directe*, chodzi o uzyskanie rozszczepienia przy równoczesnym utrzymaniu pierwotnego kierunku biegu promieni w środkowej części widma, odpowiadającej linii D sodu. Tym razem musi być spełnione równanie

$$(n'_D - 1)\varphi' - (n''_D - 1)\varphi'' = 0$$

wyrażające fakt, że całkowite odchylenie promienia D wywołane przez oba pryzmaty równa się zeru.

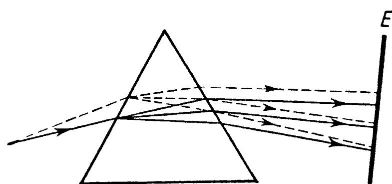
Z równania tego otrzymujemy warunek, jaki muszą spełniać kąty łamiące dwóch pryzmatów, aby ich układ był pryzmatem *à vision directe*:

$$\frac{\varphi'}{\varphi''} = \frac{n''_D - 1}{n'_D - 1}.$$

Układy *à vision directe* złożone tylko z dwóch pryzmatów dają niewielkie rozszczepienie, nawet w przypadku stosunkowo dużych kątów łamiących. Dlatego też częściej używa się układów *à vision directe* złożonych z trzech, pięciu, a nawet siedmiu pryzmatów (rys. 28.4).

28.4. Przyrządy spektralne

Do podstawowych przyrządów spektralnych należą spektroskopy (spektrografy) służące do badania i rejestracji widm oraz spektrofotometry, służące do wyznaczania natężeń linii widmowych. Rozróżniamy dwa zasadnicze typy spektroskopów: spektroskopy pryzmatyczne, oparte na zjawisku rozszczepienia, oraz spektroskopy siatkowe (dyfrakcyjne), oparte na zjawisku ugięcia światła przy przejściu przez siatkę dyfrakcyjną.



Rys. 28.5

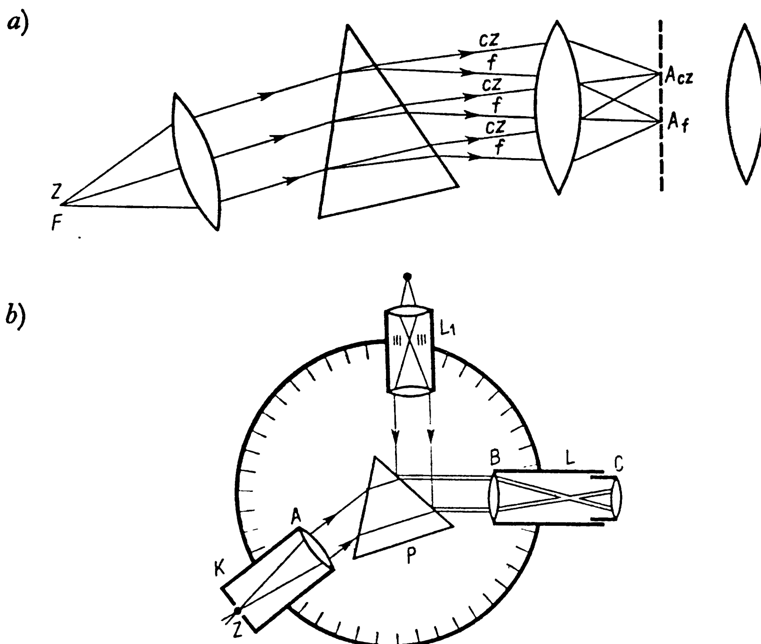
28.4.1. *Spektroskop pryzmatyczny*. Spektroskop pryzmatyczny składa się z kilku zasadniczych części, a mianowicie z kolimatora, pryzmatu i lunety. Aby dokładniej zrozumieć celowość ich stosowania, przeanalizujemy raz jeszcze załamanie światła w pryzmacie. Na rysunku 28.5 widzimy wąską wiązkę promieniowania białego, padającą bezpośrednio na pryzmat. Wobec tego, że współczynnik załamania n jest funkcją długości fali, każdy ze składników promieniowania białego załamuje się pod nieco innym kątem i na ekranie E otrzymujemy barwną wstęgę — widmo ciągłe. W widmie ciągłym występowałyby czyste barwy widmowe, tzn. do danego miejsca ekranu dochodziłoby promieniowanie o ściśle określonej długości fali jedynie w przypadku, gdyby wiązka padająca była nieskończenie wąska. W praktyce takich wiązek nie ma. Wiązka padająca jest rozciągła. Każdy z promieni wchodzących w jej skład ulega rozszczepieniu i na ekranie pokrywają się promienie różnych barw, nie dając czystego widma ciągłego.

Dogodniejszy zestaw optyczny do badania widma otrzymamy ustawiając źródło światła w ognisku soczewki zbierającej i rzucając wychodzącą z niej równoległą wiązkę na pryzmat. Promienie określonej długości fali wyjdą po załamaniu z pryzmatu jako wiązki równoległe. Na rys. 28.6a zaznaczony jest bieg równoległych promieni czerwonych i fioletowych. Ustawiając na drodze tych promieni soczewkę zbierającą achromatyczną doprowadzimy je do przecięcia się w płaszczyźnie ogniskowej soczewki. W miejscu przecięcia promieni czerwonych powstanie obraz źródła Z o barwie czerwonej. W innym miejscu powstanie obraz źródła Z o innej barwie. Utworzone obrazy można oglą-

dać w powiększeniu przez następną soczewkę ustawioną w odległości mniejszej od jej odległości ogniskowej, a więc działającej jak lupa.

W opisanym zestawie przyrządów mamy już wszystkie podstawowe elementy spektroskopu pryzmatycznego (rys. 28.6b).

Kolimator K jest to metalowy tubus zamknięty z jednej strony soczewką zbierającą A, a z drugiej strony — zasłoną metalową ze szczeliną w części środkowej. Szerokość szczeliny można dowolnie zmieniać. Tuż przed szczeliną ustawiamy badane źródło światła i traktujemy szczelinę jako źródło promieniowania Z. Długość tubusa odpowiada dokładnie odległości ogniskowej soczewki kolimatora. Dzięki temu z kolimatora wychodzi wiązka promieni równoległych i pada na pryzmat P. Po załamaniu w tym pryzmacie otrzymujemy wiązki równoległe promieni monochromatycznych, które padają na obiektyw B lunety L. Obrazy rzeczywiste szczeliny Z o różnych barwach tworzą się w płaszczyźnie ogniskowej obiektywu i mogą być oglądane przez okular C lunety jak przez lupę.



Rys. 28.6

Kolimator i luneta są umieszczone poziomo na okrągłym stoliku. Obwód stolika zaopatrzony jest w dokładną podziałkę kątową z noniuszem (nie zaznaczonym na rysunku). Kolimator jest nieruchomy, natomiast lunetę można dowolnie przesuwać.

Jeżeli źródłem promieniowania jest rozżarzone ciało stałe lub ciekłe, to przy dostatecznie wysokiej temperaturze mamy do czynienia ze światłem białym. Ustawiając takie źródło przed szczeliną kolimatora otrzymujemy po rozszczepieniu w pryzmacie nieskończenie wiele obrazów szczeliny, obrazów liniowych, dzięki jej liniowemu kształ-

towi. Obrazy te grupują się obok siebie tworząc na ekranie barwną wstęgę z ciągłym przejściem od jednej barwy do drugiej, czyli znane nam już *widmo ciągłe*.

Inaczej sprawa się przedstawia, gdy źródłem promieniowania są pobudzone do świecenia atomy lub cząsteczki ciał w postaci pary lub gazu. Takie źródło emituje tylko promieniowanie o określonych długościach fal, charakterystycznych dla danego rodzaju atomów. Wtedy w polu widzenia okularu lunety dostrzeżemy oddzielne wąskie, barwne linie. Mówimy wówczas o *widmie liniowym*, charakterystycznym dla atomów wysyłających światło. Na przykład atomy wodoru dają w dziedzinie widzialnej charakterystyczne linie widmowe o długościach fal

6568,8 Å (linia czerwona),
4861,3 Å (linia niebieska),
4340,4 Å (linia niebieska),
4101,7 Å (linia fioletowa),
3970,0 Å (linia fioletowa),
3889,0 Å (linia fioletowa).

Jeszcze inaczej wygląda widmo, gdy ciałem pobudzonym do świecenia są nie oddzielne atomy, lecz cząsteczki pierwiastków lub związków. Wtedy linie grupują się bardzo gęsto w różnych obszarach widma, tworząc tzw. pasma. Stąd pochodzi nazwa *widma pasmowe*. Do tych spraw wrócimy jeszcze w punktach 28.5.1 i 28.5.2.

Za pomocą spektroskopu wyznaczamy długości fal poszczególnych linii widmowych. W tym celu musimy spektroskop wycechować wykreślając charakterystyczną dla niego krzywą, zwaną krzywą dyspersji. Jest to krzywa zależności między długością fali i położeniem odpowiadającej jej linii w polu widzenia lunety. Do dokładnego wyznaczenia położenia danej linii widmowej potrzebna jest pewna skala odniesienia, służąca jako tło dla powstającego widma. Taką skalę pomocniczą w polu widzenia lunety otrzymujemy umieszczając na stoliku spektroskopu dodatkową lunetkę L_1 (rys. 28.6b). W płaszczyźnie ogniskowej soczewki zbierającej tej lunetki znajduje się silnie oświetlona skala. Promienie oświetlające skalę przechodzą następnie przez tę soczewkę i jako wiązka równoległa padają na ściankę boczną pryzmatu. Przy odpowiednim ustawieniu lunetki L_1 promienie odbite od ścianki pryzmatu pokrywają się z promieniami załamanymi, które wychodzą z pryzmatu. Obie wiązki promieni równoległych wytwarzają obrazy w płaszczyźnie ogniskowej obiektywu lunety L. Spoglądając przez okular widzimy obraz widma na tle obrazu skali i możemy dokładnie określać położenie każdej linii widmowej.

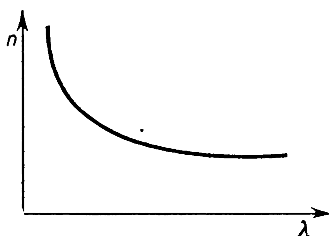
Podczas cechowania spektroskopu posługujemy się źródłami o znanych długościach fal, wyznaczonych np. za pomocą siatki dyfrakcyjnej. Na osi odciętych odkładamy długości fal, na osi rzędnych — podziałki skali pomocniczej. Otrzymana w ten sposób krzywa cechowania (rys. 28.7) jest charakterystyczna dla danego pryzmatu spektroskopu. Znając jej przebieg można wyznaczać długości fal linii widmowych dowolnych źródeł.

28.4.2. *Spektroskopy siatkowe* (*dyfrakcyjne*). Innym rodzajem przyrządu służącego do badania widma i pomiaru długości fal jest spektroskop siatkowy, którego działanie opiera się na zjawisku ugięcia światła.

Spektroskop siatkowy składa się — podobnie jak spektroskop pryzmatyczny — z podstawy kołowej zaopatrzonej w podziałkę z noniuszem, kolimatora i lunety. Na stoliku tego przyrządu umieszczona jest siatka dyfrakcyjna zamiast pryzmatu. Promienie z kolimatora padają prostopadle na siatkę dyfrakcyjną. Ustawiając oś lunety na przedłużeniu osi kolimatora otrzymamy na przecięciu nici pajęczych umieszczonych w polu widzenia okularu lunety obraz szczeliny, wytworzony przez promienie równoległe, nieugięte na siatce. Przesuwając lunetę od tego położenia w lewo lub w prawo otrzymujemy na przecięciu nici pajęczych kolejno widma pierwszego, drugiego, ... m-tego rzędu danej barwy. Jeśli źródłem promieniowania jest para lub gaz, otrzymujemy widmo liniowe. Odczytując na podziałce spektrometru kąt między dwoma ustawieniami lunety, odpowiadającymi symetrycznym położeniom tej samej linii widmowej względem prążka centralnego, znajdujemy podwojony kąt ugięcia związany z danym rzędem widma. Ze wzoru dla siatki dyfrakcyjnej

$$d \sin \alpha = m\lambda$$

znajdujemy długość fali.

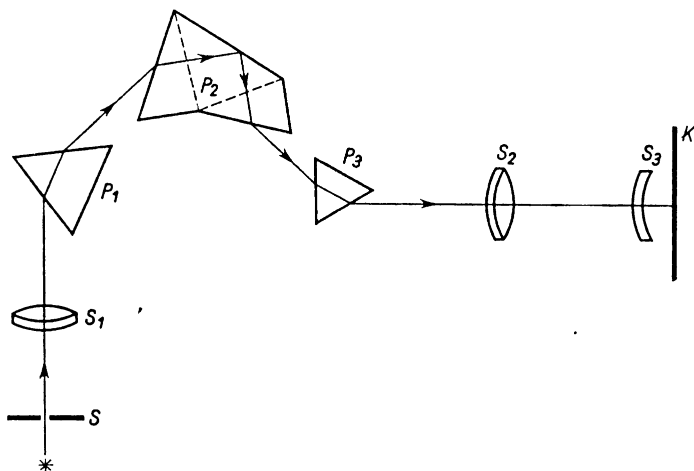


Rys. 28.7

28.4.3. *Spektrografy.* Umieszczając w płaszczyźnie ogniskowej obiektywu lunety L spektroskopu pryzmatycznego lub siatkowego kliszę fotograficzną, zastępujemy obserwację wizualną widma obserwacją fotograficzną. W ten sposób przekształcamy spektroskop na spektrograf.

Dla przykładu omówimy szczegółowiej budowę spektrografu pryzmatycznego z optyką szklaną typu ISP 51. Na rys. 28.8 przedstawiono schemat układu optycznego tego spektrografu. Widzimy tam szczelinę S z możliwością regulowania szerokości z dokładnością do 1 μm. Dalej mamy soczewkę S_1 stanowiącą achromatyczny obiektyw kolimatora. Z kolei w układzie optycznym rozróżniamy pryzmaty P_1, P_2 i P_3. Pryzmaty P_1 i P_3 są równoramienne i mają jednakowe kąty łamiące, wynoszące 63°. Pryzmat P_2 jest pryzmatem złożonym, typu pryzmatu Abbégo. W tym ostatnim można wyróżnić 3 pryzmaty składowe: dwa boczne, o stosunkowo niewielkim kącie łamiącym, i środkowy — prostokątny, równoramienny. Na tylnej ściance tego prostokątnego pryzmatu następuje całkowite wewnętrzne odbicie promieni. Kąt między promieniem padającym na pryzmat Abbégo i wychodzącym z niego wynosi 90°. Po przejściu przez wszystkie pryzmaty promienie padają na soczewki S_2 i S_3 kamery fotograficznej i po przejściu przez nie tworzą obraz widma na kliszy K.

28.4.4. *Spektrofotometry*. Spektrofotometry są to przyrządy służące do porównywania natężeń dwóch wiązek promieni o tej samej barwie. Spektrofotometry mogą służyć do badania widm pochłaniania i widm np. fluorescencji. W pierwszym przypadku obie porównywane wiązki pochodzą z tego samego źródła — jedna z wiązek bezpośrednio ze źródła skierowana jest na układ optyczny fotometru, druga natomiast trafia na ten



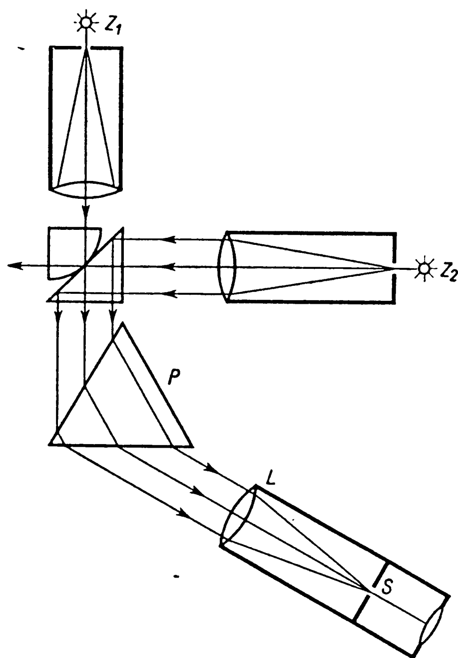
Rys. 28.8

układ po osłabieniu przez badane ciało absorbujące (płytka absorbująca o pewnej grubości x, roztwór absorbujący promieniowanie itp.). W drugim przypadku porównywane są wiązki pochodzące ze źródła badanego i ze źródła przyjętego za wzorcowe.

Nowymi elementami w spektrofotometrze — w porównaniu z fotometrem — są: układ rozszczepiający światło (a więc pryzmat, siatka dyfrakcyjna) i szczelina przepuszczająca wiązkę możliwie monochromatyczną. Rysunek 28.9 przedstawia modyfikację tego typu wprowadzoną do znanego nam już fotometru Brodhuna–Lummera. Jest to przykład spektrofotometru wizualnego. Nowością jest pryzmat P i szczelina S w płaszczyźnie ogniskowej obiektywu lunety L. Jednym z poprzednio omawianych sposobów doprowadzamy do wyrównania natężeń oświetlenia części oglądanego obrazu utworzonego przez promienie pochodzące ze źródła badanego i wzorcowego.

W większości spektrofotometrów obecnie stosowanych subiektywną, wizualną metodę obserwacji dwóch wiązek zastępują metody obiektywne. Na przykład obie porównywane monochromatyczne wiązki można skierować kolejno na komórkę fotoelektryczną (lub termoogniwo) włączoną do odpowiedniego obwodu elektrycznego. Porównanie wskazań przyrządu pomiarowego pozwala oceniać stosunki natężeń obu badanych wiązek monochromatycznych.

W innych znanych metodach spektrofotometrycznych wykorzystuje się możliwości ustalania stosunków natężeń różnych linii widmowych, sfotografowanych za pomocą spektrografu, na podstawie porównywania stopnia zaczernienia emulsji kliszy fotograficznej. Ten sam cel można też osiągnąć znacznie prościej i bardziej bezpośrednio,



Rys. 28.9

używając tzw. fotopowielaczy. Budowę i działanie fotopowielaczy omówimy szczegółowiej w § 29.7. W tej chwili potraktujmy to urządzenie jako rodzaj komórki fotoelektrycznej o silnie zwielokrotnionej czułości na padające promieniowanie. Niech taki fotopowielacz połączony z czułym galwanometrem będzie umieszczony zamiast kliszy fotograficznej w komorze fotograficznej spektrografu za wąską szczeliną. Wychylenie galwanometru można w przybliżeniu uważać za proporcjonalne do natężenia linii widmowej oświetlającej szczelinę fotopowielacza.

28.5. Podstawowe wiadomości o analizie widmowej

W ostatnich rozdziałach podkreślaliśmy kilkakrotnie możliwość otrzymywania widma i możliwość wyznaczania długości fal świetlnych. W punktach 28.4.1 i 28.4.3 mówiliśmy o budowie spektroskopów i spektrografów pryzmatycznych, umożliwiających względny pomiar długości fali, w punkcie 28.4.2 była mowa o spektroskopie siatkowym i o bezwzględnym pomiarze λ, a w punkcie 25.6.1 omawiane było doświadczenie Younga, które pozwala również na oszacowanie wartości λ. Była też mowa o spektrofotometrach pozwalających na wyznaczanie stosunku natężeń różnych linii widmowych.

Na pomiarze długości fali opiera się *analiza widmowa*, czyli określanie jakościowe i ilościowe składu chemicznego ciał na podstawie widm ich promieniowania. Wchodzi tu w grę zarówno promieniowanie widzialne, jak nadfiolet i podczerwień. Analiza wid-

mowa opiera się na fakcie, stwierdzonym w roku 1859 przez Bunsena i Kirchhoffa, że gazy i pary pobudzone do świecenia mają charakterystyczne widma.

Analiza widmowa umożliwiła odkrycie nowych pierwiastków. W ten sposób w latach sześćdziesiątych ubiegłego stulecia odkryte zostały: cez, rubid, tal, ind i gal. W latach późniejszych wykryto spektroskopowo pierwiastki ziem rzadkich i gazy szlachetne. W ostatnich dziesiątkach lat analiza widmowa jest szeroko stosowaną metodą wykonywania szybkich seryjnych analiz, mających na celu przede wszystkim wykrywanie pierwiastków śladowych. W tym ostatnim przypadku wyniki osiągane na drodze analizy spektralnej są nawet bardziej dokładne niż wyniki analiz chemicznych. Analiza widmowa jest nieocenioną metodą badawczą w przypadkach, gdy rozporządzamy małymi ilościami ciała badanego. Zyskuje ona coraz poważniejsze znaczenie w nowoczesnej metalurgii, zarówno przy badaniu zanieczyszczeń tzw. metali chemicznie czystych, jak i przy badaniu stopów, gdyż występowanie minimalnych domieszek, łatwych do wykrycia na drodze spektralnej, może w sposób zasadniczy wpływać na właściwości badanych substancji. Warto też wspomnieć o stale rosnącym jej zastosowaniu w badaniach geochemicznych, mineralogicznych, biochemicznych i astrofizycznych. Interesujące jest też zastosowanie analizy spektralnej do badań przedmiotów muzealnych. Ustalenie spektrograficzne składu farb pozwala na identyfikację oryginalnych obrazów i polichromii, podobnie badanie wyrobów ceramicznych i metalowych pozwala na wykrywanie charakterystycznych minimalnych domieszek, co ułatwia ustalenie pochodzenia i miejsca wykonania zabytków muzealnych wspomnianego typu. Badania spektroskopowe pozwoliły między innymi na rozwiązanie tajemnicy sztuki balsamowania zwłok, stosowanej przez kapłanów starożytnego Egiptu.

Dużą wykrywalnością spektralną odznaczają się pierwiastki metaliczne, stanowiące około $\frac{3}{4}$ ogólnej liczby znanych pierwiastków. Znane też są i stosowane w praktyce, ale bardziej kłopotliwe i mniej czułe, metody oznaczania spektralnego pierwiastków niemetalicznych, takich jak siarka, fosfor, chlorowce itp.

28.5.1. *Widma emisyjne.* Wiemy już, że podstawowymi przyrządami służącymi do otrzymywania widm są spektroskopy i spektrografy pryzmatyczne lub siatkowe.

Jeśli przed szczeliną kolimatora umieścimy źródło promieniowania w postaci świecącego ciała ciekłego lub rozżarzonego ciała stałego, to otrzymujemy widmo ciągłe. Promieniowanie włókna żarówki, promieniowanie Słońca i promieniowanie lampy łukowej (temperatura około 4000°C na węglu dodatnim) daje widmo ciągłe. Widmo ciągłe nie może być wykorzystywane do celów analizy widmowej emisyjnej.

Gdy źródłem promieniowania są atomy lub cząsteczki, spektrometr daje widmo nieciągłe, tzw. widmo prążkowe, składające się z oddzielnych wąskich linii lub też z linii grupujących się w pasma. *Widma* pierwsze, *liniowe*, charakterystyczne są dla atomów i jonów emitujących światło, *widma* drugie, *pasmowe*, powstają przy promieniowaniu cząsteczek. Inne prawa wiążą oddzielne linie w serie w przypadku widm liniowych (por. § 30.1), a inne odpowiadają ich grupowaniu się w pasma. W specjalnych atlasach linii widmowych podane są widma poszczególnych ciał. Przykład widma liniowego przedstawiony jest na rys. 28.10.

W analizie widmowej wykorzystywane są zarówno widma liniowe, jak i pasmowe.

Wzbudzanie świecenia może się odbywać w sposób bardzo prosty, np. przez wprowadzenie substancji łatwo parującej do płomienia palnika Bunsena. W ten sposób otrzymuje się widma lotnych metali, takich jak sód, potas, lit, wapń, stront, bar. Można też wprowadzać do płomienia sole łatwo dysocjujące w wysokiej temperaturze. Wprowadzenie np. chlorku sodu na pręciku platynowym lub na kawałku azbestu powoduje wystąpienie charakterystycznego dla sodu żółtego zabarwienia płomienia, jak również pojawienie się w widmie charakterystycznych linii żółtych sodu w dziedzinie widzialnej. Inne metody wzbudzania widm sprowadzają się do wykorzystywania łuku elektrycznego lub iskry elektrycznej. Gazy rozrzedzone można pobudzać do świecenia stosując rurki geisslerowskie.

Analiza widmowa jest metodą badawczą bardzo czułą, pozwala na przykład na wykrycie sodu w ilości $3 \cdot 10^{-7}$ mg. Z natężenia linii widmowych można wnioskować o stężeniu danego pierwiastka w badanej próbce.

28.5.2. *Widma absorpcyjne.* Na wstępie poświęcimy nieco uwagi promieniowaniu dochodzącemu do powierzchni Ziemi od Słońca. Promieniowanie to pochodzi od zewnętrznej powierzchni Słońca i daje widmo ciągłe wskutek silnego zgęszczenia materii na Słońcu, wynikającego z wielkich ciśnień tam panujących. Maksimum natężenia promieniowania przypada na dziedzinę zieloną części widzialnej widma ($\lambda \approx 5,5 \times \times 10^{-7}$ m). Aczkolwiek Słońce wysyła stosunkowo dużo nadfioletu, do powierzchni Ziemi dochodzi go jednak stosunkowo niewiele. Jedną z przyczyn tego stanu rzeczy jest fakt, że pod wpływem krótkofalowego nadfioletu w górnych warstwach atmosfery wytwarza się ozon, który ma właściwość pochłaniania nadfioletu krótszego od $3 \cdot 10^{-7}$ m. Największa ilość ozonu występuje na wysokości około 23 km nad powierzchnią Ziemi. Ogólna jego ilość jest jednak niewielka. Gdyby cała ta ilość ozonu wytworzyła warstwę o ciśnieniu jednej atmosfery, to grubość warstwy ozonu otaczającej Ziemię byłaby rzędu 1 mm. Ta stosunkowo bardzo niewielka ilość ozonu pochłania całkowicie promieniowanie krótsze od $3 \cdot 10^{-7}$ m. Nadfiolet bliższy, o długościach fal większych od $3 \cdot 10^{-7}$ m, dociera do powierzchni Ziemi również w niewielkiej ilości, gdyż ulega pochłanianiu przez atmosferę ziemską. W górach działanie opalające Słońca zawdzięczamy większemu niż w dolinach natężeniu promieniowania nadfioletowego.

Dokładne badanie widma słonecznego w dziedzinie widzialnej wykazało istnienie ciemnych linii na tle widma ciągłego (rys. 28.11). Linie te powstają również wskutek absorpcji promieniowania. Linie absorpcyjne w widmie słonecznym badane były przez Fraunhofera i dlatego nadano im nazwę linii Fraunhofera. Mówiliśmy już o nich w § 28.2.

Aby zrozumieć mechanizm powstawania linii absorpcyjnych, przeprowadzamy następujące doświadczenie. Badamy światło lampy łukowej za pomocą spektroskopu. Otrzymujemy widmo emisyjne ciągłe, gdyż źródłem promieniowania są rozżarzone węgle. Jeśli jednak na drodze wiązki światła biegnącej od lampy łukowej ustawimy płomień gazowy, do którego wprowadzimy sód, to na tle widma ciągłego pojawi się ciemny prążek w tym miejscu, które odpowiadałoby wystąpieniu żółtej linii emisyjnej sodu. Ściślej mówiąc, pojawią się dwa ciemne prążki w tych miejscach, które odpowiadałyby powstaniu żółtych linii emisyjnych sodu o długościach fal 5890 Å i 5896 Å. Ponieważ

różnica długości fal jest mała, obie linie widzimy oddzielnie tylko przy użyciu spektroskopów o dużej zdolności rozdzielczej. W przeciwnym przypadku zlewają się one w jeden prążek: żółty — przy badaniu emisji, ciemny — przy badaniu absorpcji. Pojawienie się ciemnego absorpcyjnego prążka sodu tłumaczy się tym, że atomy sodu wprowadzone do płomienia gazowego pochłaniają z promieniowania lampy łukowej ten rodzaj promieniowania, które same zdolne są wysyłać. Równocześnie wysyłają one ten sam rodzaj promieniowania. Ponieważ jednak promieniowanie to rozchodzi się we wszystkich kierunkach, więc do spektroskopu wpada tylko mały procent emitowanej energii. W związku z tym — przez kontrast z jasnym tłem widma ciągłego lampy łukowej — widzimy na ekranie ciemny prążek sodu.

Podobnie tłumaczy się powstawanie linii absorpcyjnych Słońca. Powierzchnia Słońca o temperaturze około 6000°C wysyła promieniowanie dające widmo ciągłe. Promieniowanie to przechodzi przez atmosfery gazowe Słońca i Ziemi, mające niższe temperatury. Różne składniki gazowe tych atmosfer pochłaniają charakterystyczne dla siebie długości fal, wysyłając ze swej strony promieniowanie o tej samej długości. Jak już wiemy, takie promieniowanie zaznacza się w widmie w postaci ciemnych linii na jasnym tle widma ciągłego. Rysunek 28.11 przedstawia rozmieszczenie kilku bardziej intensywnych linii Fraunhofera w skali długości fal. W dolnej części rysunku podane są symbole pierwiastków absorbujących, a w górnej — długości fal poszczególnych linii. Jak widać, poszczególne linie absorpcyjne widma słonecznego zostały przyporządkowane pierwiastkom znanym na Ziemi. Mamy tam m.in. wapń, sód, wodór, tlen. Potwierdza to, jak zresztą i inne badania, jedność budowy wszechświata. Interesujący jest fakt, że paru linii absorpcyjnych Słońca nie udało się początkowo przypisać żadnemu z pierwiastków znanych na Ziemi. Wysunięto hipotezę, że są to linie pierwiastka występującego jedynie na Słońcu. Pierwiastek ten nazwano helem. Po pewnym czasie (po 26 latach), w r. 1895, Ramsay wykrył na Ziemi pierwiastek o takich właśnie liniach emisyjnych.

Widma absorpcyjne mogą mieć nie tylko charakter liniowy, ale też ciągły i pasmowy. Przy przejściu promieniowania przez liczne ciała stałe i ciekłe, a nawet przez gazy o dużym ciśnieniu, obserwujemy w widmie wygaszenia lub silne osłabienia promieniowania nieraz w szerokim pasmie, czyli w szerokim zakresie długości fal.

28.6. Barwy ciał. Filtry absorpcyjne i interferencyjne

W paragrafie tym dość dużo uwagi poświęcimy zagadnieniom związanym z barwą ciał. Przeanalizujmy to zagadnienie nieco dokładniej.

Omawiając zagadnienie barw ciał musimy uwzględnić cztery czynniki, a mianowicie: właściwości oka ludzkiego, rodzaj powierzchni ciała rozpraszającego, rodzaj źródła światła oraz ośrodek, przez który promieniowanie przechodzi.

Oko ludzkie jest przyrządem rejestrującym dochodzące promieniowanie. Jak wiemy, oko jest wrażliwe na promieniowanie zawarte w przedziale długości fal od $4 \cdot 10^{-7}$ m do $8 \cdot 10^{-7}$ m. Gdy do oka dochodzi wiązka promieniowania o pewnej określonej długości fali z wymienionego wyżej zakresu, wówczas odbieramy wrażenie barwne, od-

powiadające czystej barwie widmowej i mamy do czynienia z wiązką światła monochromatycznego.

Pamiętamy z rozdziału poświęconego budowie oka, że wrażliwość na barwy wiąże się z istnieniem na siatkówce czopków. Według teorii Younga–Helmholtza czopki mają trzy rodzaje substancji wrażliwych na trzy podstawowe barwy: czerwoną, zieloną i niebieską. Oko nie jest jednak niezawodnym przyrządem służącym do określania barw promieniowania widzialnego. Ludzie o normalnej reakcji na barwy zaliczani są do tzw. *trichromatów*. Daltoniści to najczęściej dichromaci reagujący tylko na dwie podstawowe barwy, gdyż brak im jednej z substancji wrażliwych na barwy. Daltonizm jest wadą stosunkowo rozpowszechnioną, szczególnie u mężczyzn; około 30% mężczyzn wykazuje tę wadę wzroku. *Monochromaci* to ludzie zupełnie niewrażliwi na barwy, reagujący tylko na stopnie jasności (szarości).

Gdy do oka dochodzi promieniowanie widzialne o różnych długościach fali i różnych natężeniach, oko sumuje wszystkie podniety i człowiek odbiera pewne wrażenie barwne. Nie jest to już jednak czysta barwa widmowa. Mówimy wtedy o *addytywnym powstawaniu barw*. Promieniowanie złożone o takim składzie i takich stosunkach natężeń, jakie występują w promieniowaniu słonecznym, wywołuje wrażenie światła białego. Wystarczy usunąć którąkolwiek z długości fal lub zmienić stosunek natężeń, by zamiast światła białego otrzymać wrażenie barwne. Sposób *wywoływania barw* przez usuwanie pewnych składników padającego promieniowania nazywamy *substraktywnym*. Na przykład usunięcie z promieniowania białego promieni fioletowych spowoduje wystąpienie barwy żółtozielonej, usunięcie promieniowania czerwonego wywoła wrażenie barwy zielononiebieskiej. Takie pary czy trójki barw, które w sumie dają wrażenie światła białego nazywamy barwami dopełniającymi. Zestawienie kilku par barw dopełniających podane jest w poniższej tabeli.

Barwy dopełniające

czerwona	6560 Å	↔	zielononiebieska	4920 Å
pomarańczowa	6070 Å	↔	niebieska	4900 Å
żółta	5740 Å	↔	niebieska	4820 Å
zielonożółta	5640 Å	↔	fioletowa	4320 Å

Obserwowane w świetle dziennym przedmioty barwne są to najczęściej ciała, które podczas *procesu rozpraszania* absorbują na swej powierzchni część promieniowania widzialnego. Na przykład czerwony płatek róży, dzięki cząsteczkom barwnika zawartym w płatku, zaabsorbował wszystkie składniki promieniowania padającego, z wyjątkiem czerwonego. Rozproszenie promieni czerwonych decyduje o odbieranym wrażeniu czerwieni. Zielone części roślin zawdzięczają swoje zabarwienie zawartości chlorofilu. Chlorofil ma pasma absorpcyjne w dziedzinie czerwonej, niebieskiej i fioletowej. Mieszanina pozostałych składników rozproszonego światła białego daje wrażenie zieleni.

Absorpcja, charakteryzująca szereg ciał barwnych przezroczystych, decyduje o możli-

wości stosowania ich jako filtrów tzw. absorpcyjnych. Wymienimy dla przykładu różnobarwne filtry fotograficzne, umiejętnie przez fotografa dobierane do warunków oświetleniowych, filtry spektralne (np. filtr Wooda — zatrzymujący promienie widzialne, a przepuszczający nadfiolet), filtry stosowane jako okulary dla spawaczy — absorbujące nadfiolet i podczerwień itd.

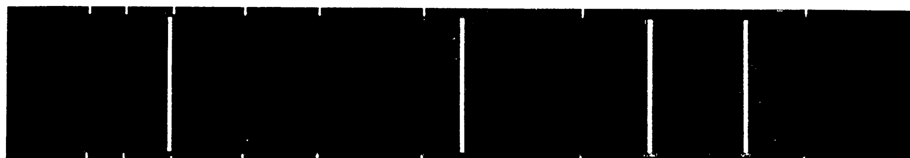
Ciekawą przyczyną występowania barw jest tzw. *selektywne odbicie*. Niektóre ciała przepuszczają pewne barwy, inne absorbują, a równocześnie silnie odbijają. Na przykład plamka z wyschniętego czerwonego atramentu na szkle jest w świetle przepuszczonym czerwona, natomiast obserwowana z boku w świetle odbitym ma wyraźne zabarwienie zielone. Podobne selektywne odbicie wykazują metale i temu właśnie zjawisku zawdzięczają swą barwę np. miedź i złoto. Złoto odbija selektywnie promienie czerwonożółte. Cieniutka folia złota obserwowana w świetle przepuszczonym wykazuje zabarwienie zielonkawe.

Interesująco przedstawiają się barwne tkaniny oglądane w świetle zmieniającym barwę, np. tkanina czerwona w świetle białym zachowuje swą barwę przy zastosowaniu oświetlenia czerwonego, a wydaje się czarna, gdy zastosujemy inne monochromatyczne źródło światła. Tkaniny zielone charakteryzuje bardziej złożona absorpcja i w związku z tym przyjmują one odcienie zielone, żółtozielone lub niebieskawe w zależności od oświetlenia.

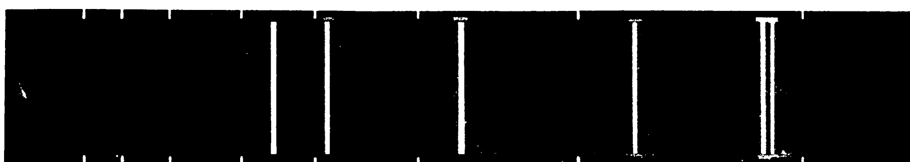
Inna jest przyczyna występowania czerwonego zabarwienia tarczy zachodzącego lub wschodzącego Słońca. W tym przypadku decydującą rolę odgrywa ośrodek, przez który promieniowanie od źródła dochodzi do naszego oka. W godzinach rannych i wieczornych, w związku z niskim położeniem Słońca nad horyzontem, światło słoneczne przechodzi przez grubą warstwę atmosfery ziemskiej. Po drodze światło ulega rozproszeniu na różnych zawiesinach (kurz), a nawet na przypadkowych zgęszczeniach cząsteczek gazu wynikających z ich ruchów cieplnych. W większym stopniu rozproszeniu ulegają fale krótkie, a więc promienie fioletowe, niebieskie i zielone. W promieniowaniu docierającym do oka obserwatora zmieniony został stosunek natężeń, charakterystyczny dla światła białego: osłabiona została część krótkofalowa i oko odbiera wrażenie barwy dopełniającej, widzi czerwoną tarczę Słońca. W południe, gdy kierunek padania promieniowania słonecznego na powierzchnię Ziemi jest zbliżony do prostopadłego, wpływ atmosfery ziemskiej jest nieznaczny.

Skutki rozpraszającego działania atmosfery obserwujemy też na przykładzie błękitu nieba. Promieniowanie słoneczne przechodząc przez górne warstwy atmosfery ulega rozproszeniu. Smoluchowski wykazał, że przyczyna rozproszenia tkwi we fluktuacjach gęstości powietrza. Promieniowanie krótkofalowe fioletowe i niebieskie ulega rozproszeniu w silniejszym stopniu niż promieniowanie czerwone. Promienie fioletowe i niebieskie, rozproszone w górnych warstwach atmosfery, trafiają do oka ludzkiego. Dzięki większej wrażliwości oka na promienie niebieskie, odbieramy wrażenie błękitu.

Wpływ wielkości cząsteczek zawiesin na barwę światła rozproszonego można wykazać doświadczalnie biorąc np. roztwór tiosiarczanu sodu i dodając do niego rozcieńczony kwas siarkowy. Wytrącająca się z roztworu siarka występuje w ziarnach, które stopniowo powiększają swe rozmiary. Równocześnie zmienia się barwa roztworu w świet-



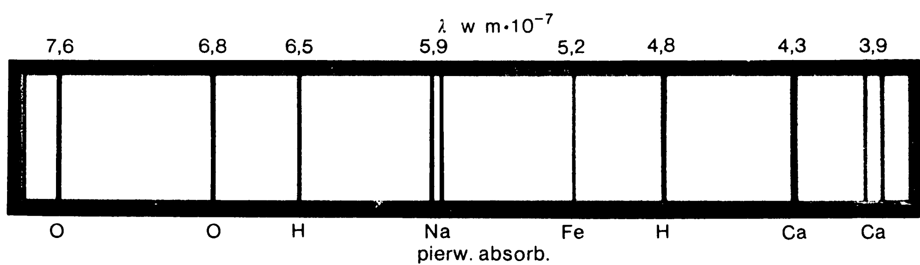
wodór



rtęć

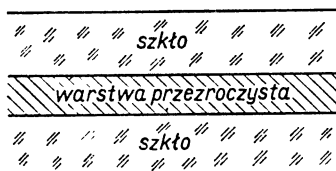


neon



le rozproszonym — od fioletu do czerwieni. Koloidalny roztwór złota daje w świetle rozproszonym piękny odcień purpury.

Prócz barw związanych z omówionymi wyżej czynnikami mogą jeszcze występować *barwy interferencyjne*. Barwy te, jak sama nazwa wskazuje, związane są ze zjawiskiem interferencji, która może mieć miejsce wtedy, gdy promieniowanie spójne dochodzące od źródła do naszego oka biegnie po różnych drogach. Różnicy dróg promieni odpowiada różnica faz. Jeśli dana różnica faz utrzymuje się przez pewien czas — wyniki interferencji można dostrzec. Przy spotkaniu promieni w fazach przeciwnych następuje wygaszenie promieniowania danej długości fali. Ta sama różnica dróg optycznych może prowadzić do spotkania w fazach zgodnych promieniowania o innej długości fali i wtedy w wyniku interferencji występuje wzmocnienie tego drugiego rodzaju promieniowania. A zatem ulega zmianie rozkład natężeń odbieranego światła: jedne składniki zostaną wygaszone, inne wzmocnione. W wyniku końcowym otrzymujemy wrażenie barwne. Mówiliśmy o tym szeroko w § 25.3. Omawiane tam zjawiska interferencji promieni znalazły zastosowanie w tzw. filtrach interferencyjnych. Filtr taki składa się z 2 płasko-równoległych płytek szklanych (rys. 28.12). Powierzchnie tych płytek zwrócone do siebie pokryte są cienką, półprzezroczystą warstwą srebra. Między tymi powierzchniami znajduje się cienka warstwa z przezroczystego materiału. Obserwujemy wyniki interferencji promieni przechodzących bezpośrednio przez filtr i promieni ulegających odbiciu od metalicznych powierzchni (por. § 25.3). Jeśli np. w przypadku prostopadłego padania ($\beta = 0$) promieni na filtr otrzymujemy w świetle przepuszczonym zabarwienie niebieskozielone, tzn. że wygaszeniu ulega część czerwona widma. Gdy będziemy powiększali kąt padania, będzie też rósł kąt załamania β, a $\cos\beta$ będzie malał i kolejno wygaszane będą fale coraz krótsze: pomarańczowa, żółta, zielona itd. Barwa światła przepuszczonego przez taki filtr będzie się zmieniała od niebieskozielonej, przez niebieską, fioletową, purpurową, pomarańczową do żółtej.



Rys. 28.12

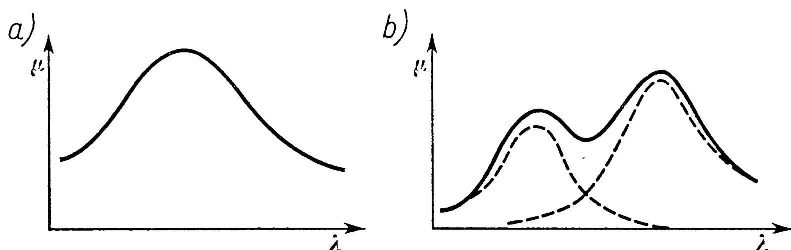
28.7. Podstawy fotokolorymetrii

Fotokolorymetria zajmuje się ustalaniem stężenia roztworu na podstawie jego barwy. Zanim przejdziemy do wyjaśnienia metod takiego pomiaru, omówimy prawo, na którym się one opierają. Jest to *prawo Beera*, obowiązujące dla roztworów o niezbyt dużym stężeniu. Według tego prawa *natężenie I światła monochromatycznego przechodzącego przez grubość x warstwy roztworu o stężeniu C maleje wykładniczo*:

$$I = I_0 e^{-\mu C x},$$

gdzie I_0 oznacza natężenie światła padającego na warstwę, a μ — współczynnik charakterystyczny dla ciała rozpuszczonego, w dość szerokich granicach niezależny od stężenia roztworu. Iloczyn $\mu C = k$ oznacza tzw. *współczynnik pochłaniania*.

Badania wykazały, że wartość współczynnika μ danego ciała rozpuszczonego jest zależna od długości fali padającego promieniowania. Zależność ta dla różnych ciał może być bardzo różna. W przypadku cieczy przebiegi $\mu = f(\lambda)$ przedstawione są przykładowo na rys. 28.13a i b. Na drugim rysunku widoczne jest nałożenie się dwóch obszarów absorpcji (linie przerywane) powodujące dość skomplikowany przebieg krzywej wypadkowej. Trzeba też podkreślić, że krzywe przedstawiające $\mu = f(\lambda)$ mogą mieć dla danej substancji rozpuszczonej przebieg zupełnie różny, gdy zmienimy rozpuszczalnik (często stwierdza się przesunięcie maksimum pochłaniania w stronę długich fal, jeśli zmianie rozpuszczalnika towarzyszy wzrost współczynnika załamania). Zmiana typu cząsteczki w roztworze (a więc np. zmiana stopnia dysocjacji, tworzące się asocjaty itp.) wpływa również na przebieg krzywych pochłaniania i powoduje duże odstępstwa od prawa Beera.



Rys. 28.13

Kolorymetryczna metoda wyznaczania stężenia może być stosowana tylko wtedy, gdy odstępstwa od prawa Beera są małe. Metoda polega na tym, że dobieramy grubości warstw roztworu badanego i roztworu o znanym stężeniu tak, aby barwy obu roztworów w świetle przechodzącym były jednakowe. Barwy te wydają się nam jednakowe wtedy, gdy wszystkie składniki w wiązkach przechodzących ulegają temu samemu osłabieniu, a więc gdy zgodnie z prawem Beera

$$e^{-\mu C_1 x_1} = e^{-\mu C_2 x_2},$$

stąd zaś otrzymujemy

$$C_1 x_1 = C_2 x_2,$$

$$\frac{C_1}{C_2} = \frac{x_2}{x_1}.$$

Pomiar staje się dokładniejszy, jeśli przez zastosowanie odpowiednio dobranych filtrów usuniemy z wiązek padających te składniki, które nie ulegają pochłanianiu w roztworach.

Kolorymetry różnego typu znalazły szerokie zastosowanie w laboratoriach naukowych, przemysłowych i medycznych.

Pytania i zadania

1. W jakich granicach długości fal zawiera się podczerwień, promieniowanie widzialne i nadfiolet?
2. Jakie są metody wzbudzania i wykrywania podczerwieni i nadfioletu?
3. W jaki sposób można otrzymywać układy achromatyczne i nie odchylające pryzmatów?
4. Za pomocą jakich wielkości można scharakteryzować dyspersję?
5. Jakie są podstawowe przyrządy spektralne? Podać zasady ich budowy i działania.
6. Na czym polega analiza widmowa? Na badaniu jakich rodzajów widm opiera się ta metoda badawcza?
7. Omówić różnorodne przyczyny powstawania barw.
8. Wymienić przykłady filtrów absorpcyjnych i interferencyjnych, wyjaśnić ich działanie i podać przykłady zastosowań.
9. Jaki jest cel badań fotokolorymetrycznych? Na jakim prawie te badania się opieraja? Czy nie ma żadnych ograniczeń stosowalności tego prawa?
10. Jakie wnioski można wyciągnąć z tego, że w widmie pewnej mgławicy stwierdzono występowanie linii wodoru H_α, ale o długości fali przesuniętej o 20 Å w stronę czerwieni?
11. Wyrazić dyspersję kątową pryzmatu o kącie łamiącym φ w zależności od $dn/d\lambda$ dla obszaru minimum odchylenia.

Odp. $$\frac{d\delta}{d\lambda} = \frac{2\sin\frac{\varphi}{2}}{\cos\frac{\delta_{\min}+\varphi}{2}}\,\frac{dn}{d\lambda}.$$

CZĘŚĆ V

Wybrane zagadnienia z fizyki współczesnej

ROZDZIAŁ 29

Podstawy fizyki współczesnej

29.1. Promieniowanie termiczne. Prawo Stefana–Boltzmanna

We wstępie do optyki mówiliśmy o rozwoju teorii promieniowania, wymieniając teorię korpuskularną Newtona, falową Huygensa i wreszcie teorię falową elektromagnetyczną Maxwella. Powróćmy obecnie do przełomu XIX i XX wieku, kiedy to — na skutek wyłaniających się przy tłumaczeniu pewnych zjawisk trudności nie do przezwyciężenia z punktu widzenia teorii fizyki klasycznej — zostają przez Plancka położone podwaliny pod nową, kwantową teorię promieniowania.

Trudności, o których wyżej mowa, wiążą się m.in. z promieniowaniem termicznym (temperaturowym). Poświęcimy temu promieniowaniu nieco więcej uwagi.

Jest rzeczą ogólnie znaną, że temperatura ciała ma wpływ na wysyłane przez nie promieniowanie. Promieniowanie widzialne pojawia się wtedy, gdy źródło ma temperaturę wyższą od 950K: w czasie ogrzewania powyżej 950K początkowo rozszerza się stopniowo część czerwona widma, w temperaturze bliskiej 1500K w widmie występuje część czerwona, żółta i zielona, a w temperaturze około 1800K widmo obejmuje już cały zakres widzialny.

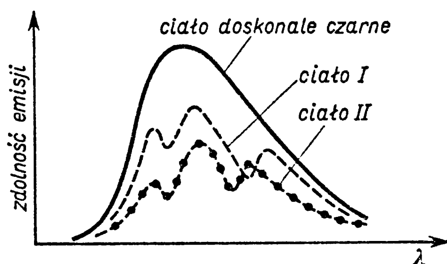
Umiejętny dobór przyrządów rozszczepiających promieniowanie (spektrometrów) pozwala stwierdzić, że widmo rozciąga się także na podczerwień (IR) i nadfiolet (UV)*, przy czym podczerwień jest głównie reprezentowana w temperaturach stosunkowo niskich (< 950K), a nadfiolet — w temperaturach wysokich.

Wyniki badań promieniowania termicznego można przedstawić na wykresie, odkładając na osi odciętych długość lub częstotliwość fali, a na osi rzędnych — tzw. *zdolność emisji*, tzn. ilość energii promienistej (odpowiadającej wąskiemu zakresowi długości fali między λ i $\lambda+d\lambda$) wysyłanej w jednostce czasu przez jednostkę powierzchni źródła, a zatem wyrażanej w W/m^2. Wykresy mogą się odnosić: a) do różnych źródeł bada-

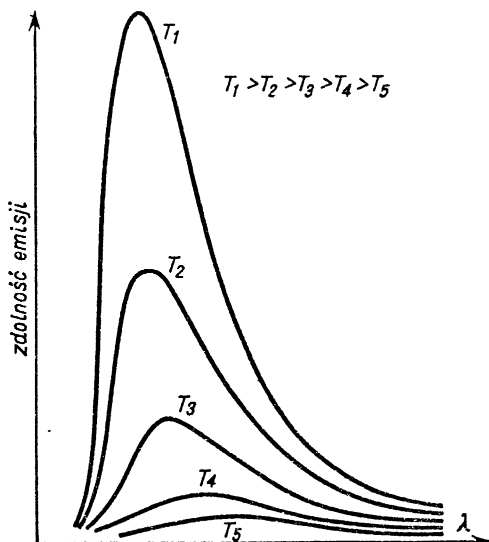
* Skróty pochodzą od nazw angielskich

nych w tej samej temperaturze (rys. 29.1), b) do tego samego źródła badanego w różnych temperaturach (rys. 29.2).

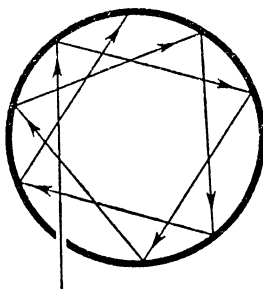
Na rysunku 29.1 przedstawione są trzy krzywe. Krzywa kreskowana i kropkowana dotyczą dwóch różnych ciał o tej samej temperaturze. Aczkolwiek poszczególnym długościom fali odpowiadają nieco różne zdolności emisji, to jednak ogólny charakter obu krzywych jest zbliżony. Z wykresów tych widać, że rodzaj źródła ma wpływ na przebieg emisji, co oczywiście utrudnia wyciąganie jakichkolwiek ogólnych wniosków. Okazało się, że można sztucznie wytworzyć źródło o maksymalnej zdolności emisji w każdej temperaturze i o 100-procentowej absorpcji padającego promieniowania również w każdej temperaturze. Ciało takie nazwano *doskonale czarnym*. Właściwości ciała doskonale czarnego wykazuje niewielki otwór utworzony w powierzchni kuli wydrążonej, poczernionej w środku. Energia promieniowania wnikająca przez taki otwór do wnętrza kuli zostaje praktycznie całkowicie pochłonięta podczas licznych odbić od powierzchni wewnętrznej (rys. 29.3). Z padającego pierwotnie na otwór promieniowania żaden promień nie przenika z kuli na zewnątrz, otwór zachowuje się więc jak ciało doskonale czarne. Wykreślając dla takiego otworu, traktowanego jako źródło, zależność zdolności emisji od długości fali otrzymamy krzywą o przebiegu takim, jaki ma krzywa ciągła



Rys. 29.1



Rys. 29.2


Rys. 29.3

na rys. 29.1. Krzywa ta jest już niezależna od charakteru źródła, a zależna tylko od temperatury. Warto od razu podkreślić, że całkowita energia promieniowania w zakresie częstotliwości od 0 do ∞ wysłana przez jednostkę powierzchni badanego źródła w jednostce czasu jest proporcjonalna do pola znajdującego się pod krzywą. Według ustalonego doświadczalnie prawa Stefana, uzasadnionego następnie teoretycznie przez Boltzmanna, *całkowita energia promieniowania widzialnego i niewidzialnego wysyłana przez jednostkę powierzchni ciała doskonale czarnego w jednostce czasu wyraża się wzorem*

$$E = \sigma T^4, \tag{29.1}$$

gdzie σ oznacza stałą Stefana–Boltzmanna, równą $5{,}669 \cdot 10^{-8} \frac{\mathrm{W}}{\mathrm{m}^2 \cdot \mathrm{K}^4}$.

29.2. Prawo Wiena

Powróćmy do wspomnianego w poprzednim paragrafie drugiego sposobu przedstawiania wyników badań promieniowania termicznego: niech stale tym samym badanym źródłem będzie ciało czarne, a każda krzywa wykresu (29.2) niech dotyczy innej temperatury. Jak widać, w miarę wzrostu temperatury nie tylko pole pod krzywą gwałtownie rośnie (zgodnie z prawem Stefana–Boltzmanna), lecz także maksimum krzywej przesuwa się wyraźnie w stronę fal krótkich. Wien stwierdził, że obowiązuje przy tym zależność

$$\lambda_{\max} T = \mathrm{const} = 2{,}897 \cdot 10^{-3} \mathrm{m} \cdot \mathrm{K}, \tag{29.2}$$

gdzie $\lambda_{\max}$ oznacza długość fali (w m), przy której występuje maksimum zdolności emisji w temperaturze TK. Równanie (29.2) wyraża *prawo przesunięć Wiena*.

29.3. Prawo Kirchhoffa

W przeglądzie praw rządzących promieniowaniem termicznym nie może zabraknąć prawa Kirchhoffa. Prawo to, wyprowadzone na podstawie rozważań termodynamicznych, podaje zależność między zdolnością emisji promieniowania o długości fali λ i zdolnością absorpcji tegoż promieniowania w tej samej temperaturze.

W paragrafie 29.1 podaliśmy już definicję zdolności emisji badanego ciała; oznaczamy ją przez $e(\lambda, T)$, gdyż jak wiemy, jest ona zależna od długości fali (częstotliwości promieniowania) i od temperatury. *Zdolność absorpcji* definiujemy inaczej: jest to liczba wskazująca, jaka część strumienia promieniowania padającego zostaje pochłonięta. Jest to zatem liczba niemianowana, w przypadku ciał rzeczywistych stanowiąca zawsze ułamek właściwy, a w przypadku modelowego ciała doskonale czarnego (ciało doskonale czarne, tj. ciało o 100% absorpcji dla wszystkich częstotliwości i temperatur) z definicji równa 1. Zdolność absorpcji badanego ciała oznaczamy przez $\alpha(\lambda, T)$, gdyż jest ona także funkcją długości fali i temperatury*.

---

* Może warto tu wspomnieć, że w przypadku padania energii promienistej na dowolną powierzchnię część energii ulega odbiciu, część jest absorbowana, a reszta jest przepuszczana. Wprowadzając pojęcia *zdolności odbijania* ϱ i *zdolności przepuszczania* τ i definiując je analogicznie do zdolności absorpcji, otrzymujemy związek

$$\alpha + \varrho + \tau = 1.$$

Oczywiście zarówno zdolność emisji e, jak i zdolność absorpcji α dotyczy możliwie wąskich przedziałów $\lambda+d\lambda$ (lub $\nu+d\nu$), czyli odnoszą się one do promieniowania możliwie monochromatycznego.

Zgodnie z prawem Kirchhoffa *stosunek zdolności emisji do zdolności absorpcji danego ciała nie zależy od rodzaju ciała, tzn. dla wszystkich ciał jest jednakową funkcją częstotliwości i temperatury*, czyli

$$\frac{e(\lambda, T)}{\alpha(\lambda, T)} = \varepsilon(\lambda, T). \tag{29.3}$$

Aby znaleźć sens fizyczny wielkości $\varepsilon(\lambda, T)$, napiszemy równanie (29.3) w odniesieniu do ciała doskonale czarnego wprowadzając wskaźnik zero:

$$\frac{e_0(\lambda, T)}{\alpha_0(\lambda, T)} = \varepsilon(\lambda, T).$$

Uwzględniając, że dla ciała doskonale czarnego $\alpha_0(\lambda, T) = 1$ otrzymamy

$$e_0(\lambda, T) = \varepsilon(\lambda, T),$$

Innymi słowy, występująca w prawie Kirchhoffa *wielkość $\varepsilon(\lambda, T)$ równa się zdolności emisji ciała doskonale czarnego dla długości fali λ w temperaturze T*:

$$\frac{e(\lambda, T)}{\alpha(\lambda, T)} = e_0(\lambda, T). \tag{29.4}$$

Czasem można spotkać inne ujęcie prawa Kirchhoffa, a mianowicie: *zdolność emisji danego ciała w danej temperaturze równa się jego zdolności absorpcji*. Przy takim ujęciu obowiązuje inna definicja zdolności emisji, a mianowicie przez *zdolność emisji danego ciała* rozumiemy stosunek

$$\frac{e(\lambda, T)}{\varepsilon_0(\lambda, T)}.$$

Wtedy, jak łatwo sprawdzić,

$$\frac{e(\lambda, T)}{\varepsilon_0(\lambda, T)} = \alpha(\lambda, T).$$

Ze związku tego widzimy, że *ciało doskonale czarne ma maksymalną zdolność emisji każdego z rodzajów promieniowania w każdej temperaturze T*.

Trzeba podkreślić, że *stosowalność prawa Kirchhoffa ogranicza się wyłącznie do promieniowania termicznego, tzn. powstającego kosztem energii cieplnej*. Nie obowiązuje ono w przypadkach luminescencji (chemiluminescencja, fotoluminescencja, elektroluminescencja itp.).

Ważnym wnioskiem, wynikającym z prawa Kirchhoffa jest to, że jeśli jakieś ciało w pewnej temperaturze T emituje promieniowanie o długości fali λ, to absorbuje ono także ten sam rodzaj promieniowania. Z odwróceniem tego wniosku należy być ostrożnym, absorpcja promieniowania o długości λ w temperaturze T nie musi iść w parze z emisją tegoż promieniowania, gdyż zdolność emisji $e(\lambda, T)$ może być równa zeru,

jeżeli dla danej temperatury i długości fali zdolność emisji ciała doskonale czarnego $e_0(\lambda, T)$ równa się zeru.

Potwierdzeniem prawa Kirchhoffa może być następujące doświadczenie. Do płomienia palnika Bunsena wprowadzamy kawałek kwarcu i grafitu. Okazuje się, że kwarc świeci słabo, grafit silnie, czyli zdolność emisji kwarcu jest mała w porównaniu ze zdolnością emisji grafitu w odniesieniu do światła widzialnego. Ale z drugiej strony wiadomo, że kwarc jest przezroczysty, a grafit nieprzezroczysty dla światła widzialnego, czyli zdolność absorpcji kwarcu jest mała, a grafitu duża. Tylko przy takim układzie wartości $e(\lambda, T)$ i $\alpha(\lambda, T)$ dla obu materiałów może być spełnione prawo Kirchhoffa.

P y t a n i e. Jak wytłumaczyć fakt, że kawałek czerwonego szkła w temperaturze pokojowej nie wysyła promieniowania o barwie zielonej, mimo że ten rodzaj promieniowania prawie w 100% pochłania?

Z prawa Kirchhoffa korzystamy też przy tłumaczeniu powstawania widma absorpcyjnego (por. punkt 28.5.2).

29.4. Kwantowość promieniowania. Wzór Plancka

Dla dalszego rozwoju teorii promieniowania podstawowe znaczenie miała omawiana w § 29.2 krzywa (rys. 29.2) przedstawiająca zdolność emisji ciała doskonale czarnego w funkcji długości fali. Dość uderzające jest podobieństwo jej przebiegu do przebiegu krzywej rozkładu prędkości Maxwella. Podobieństwo to skłoniło Wiena do poczynienia prób dostosowania do opisu promieniowania ciała doskonale czarnego równania analogicznego do równania Maxwella. Udało mu się ustalić wzór empiryczny z dwiema stałymi tak dobranymi, że zapewniały dobrą zgodność wartości otrzymywanych na podstawie tego wzoru z wynikami eksperymentalnymi w zakresie odpowiadającym falom krótkim. Niezgodność wyników w zakresie fal długich można było tłumaczyć tym, że równanie Wiena nie jest dobrze „dobrane", przypominamy bowiem, że równanie to nie wynikało z przesłanek teoretycznych charakteryzujących mechanizm emisji promieniowania.

Sytuacja niewątpliwie bardzo się zaostrzyła, gdy Rayleigh w wyniku ścisłych rozważań opartych na prawach fizyki klasycznej (równy podział energii na wszystkie stopnie swobody układu) doszedł do wzoru, który dawał wyniki zgodne z doświadczalnymi w dziedzinie fal długich, lecz całkowicie odmienne w dziedzinie fal krótkich. Ta niezgodność z doświadczeniem prawa wyprowadzonego na podstawie ścisłych klasycznych rozważań oceniona została jako tzw. katastrofa nadfioletowa, stało się bowiem jasne, że fizyka klasyczna nie jest w stanie wytłumaczyć ważnego faktu doświadczalnego.

Wyjście z tego poważnego impasu zawdzięcza fizyka Planckowi. Pierwszy jego krok był dość przypadkowy. Planck stwierdził, że wprowadzając pewne poprawki do wzoru Wiena i zmieniając jego stałe otrzymuje się wzór, który w granicy dla fal długich przechodzi we wzór Wiena, a w całym zakresie długości fal λ daje wyniki zgodne z doświadczeniem*.

---

* Dla umożliwienia porównania obu wzorów podajemy niżej:

wzór Plancka $$e(\lambda, T)d\lambda = \frac{2C_1}{\lambda^5}\frac{1}{e^{C_2/\lambda T}-1}d\lambda,$$

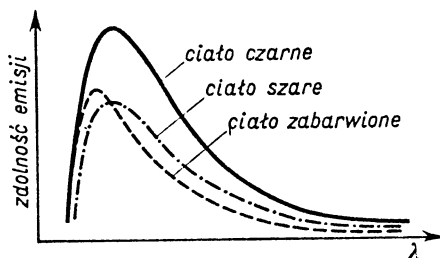
Można więc było przyjąć skorygowany wzór za słuszny. Drugim krokiem uczynionym przez Plancka było poszukiwanie poprawnej z punktu widzenia fizyki i matematyki metody uzasadnienia otrzymanego wzoru. Żadna z prób opartych na założeniach fizyki klasycznej, nie dała dobrego rezultatu. W ujęciu klasycznym zakładano zawsze, że energia może przyjmować wszelkie wartości, a zatem zmiany energii mogą zachodzić w sposób ciągły. Planck wprowadza całkowicie nowy obraz mechanizmu promieniowania. Jego zdaniem zmiany energii atomowego źródła wysyłającego promieniowanie mogą zachodzić tylko określonymi porcjami, tzn. w sposób nieciągły. Porcja wypromieniowywanej energii E (zwana obecnie *kwantem promieniowania* lub *fotonem*) wyraża się *wzorem Plancka*

$$E = h\nu, \tag{29.4}$$

gdzie h oznacza *stałą Plancka* równą $6{,}625\,3 \cdot 10^{-34}\,\mathrm{J \cdot s}$, a ν — częstotliwość emitowanego promieniowania. Cecha kwantowości promieniowania, wyrażona równaniem (29.4), okazała się niesłychanie ważna dla rozwoju fizyki współczesnej. Wspomniane równanie jest podstawą, na której rozwinęła się teoria kwantowa promieniowania. W dalszych rozdziałach tego podręcznika niejednokrotnie jeszcze wymieniać będziemy tę podstawową stałą fizyczną, jaką jest stała Plancka.

29.5. Promieniowanie termiczne ciał rzeczywistych

Większość rozważań w poprzednich paragrafach tego rozdziału poświęciliśmy ciału modelowemu — ciału doskonale czarnemu. Musimy jednak na zakończenie powrócić raz jeszcze do warunków realnych i poświęcić nieco więcej uwagi promieniowaniu termicznemu ciał rzeczywistych.



Rys. 29.4

Jak już mówiliśmy w § 29.1, krzywa zależności zdolności emisji od długości fali λ ma przebieg zależny od rodzaju ciała wysyłającego promieniowanie termiczne i odbiegający mniej lub bardziej od krzywej charakterystycznej dla ciała doskonale czarnego.

---

wzór Wiena $$e(\lambda, T)d\lambda = \frac{C'}{\lambda^5} e^{-C''/\lambda T} d\lambda,$$

gdzie C_1, C_2, C', C'' są wielkościami stałymi.

P y t a n i e. W jaki sposób ze wzoru Plancka można przejść: a) do wzoru Stefana–Boltzmanna, b) do wzoru Wiena?

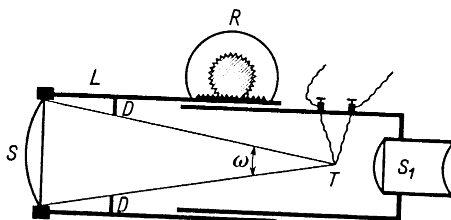
Istnieje grupa ciał tzw. szarych, dla których kształt omawianej krzywej jest analogiczny do kształtu krzywej ciała doskonale czarnego, tylko wszystkie rzędne są w tym samym stosunku zmniejszone. Taki przebieg krzywej świadczy również o tym, że zdolność absorpcji badanego ciała nie zależy od długości fali.

Dla pozostałych ciał — ciał zabarwionych — zdolność absorpcji różnie się zmienia z długością fali, a zatem omawiane krzywe, aczkolwiek zawsze leżą poniżej krzywej dla ciała czarnego o tej samej temperaturze, mają różny przebieg. Przykłady przebiegu krzywych dla ciała szarego (węgiel) i zabarwionego (metal rozżarzony) przedstawione są na rys. 29.4. Tego rodzaju ciągłe przebiegi krzywych są oczywiście charakterystyczne dla ciał stałych, gdyż jak wiadomo, gazy i pary ogrzane do wysokiej temperatury wysyłają promieniowanie o widmie prążkowym.

29.6. Pirometria optyczna

Nazwą *pirometrów* obejmujemy urządzenia służące do pomiaru wysokich temperatur. O niektórych typach pirometrów mówiliśmy już w rozdziale poświęconym termometrii (pirometry elektryczne). Obecnie zajmiemy się zasadą budowy i działania pirometrów optycznych, które znalazły szerokie zastosowanie m.in. w metalurgii do pomiarów temperatur wnętrz pieców.

Zgodnie z tym, co było powiedziane wyżej, otwór pieca można w przybliżeniu traktować jako ciało doskonale czarne. Można zatem temperaturę pieca wyznaczyć na podstawie prawa Stefana–Boltzmanna lub prawa Wiena.



Rys. 29.5

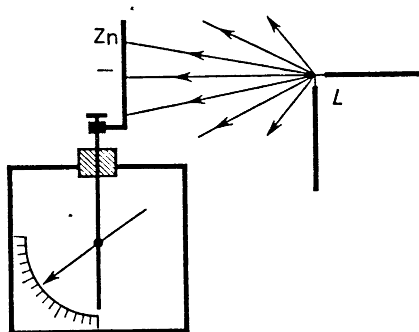
Rozważmy bliżej działanie pirometru Féry'ego, którego schemat budowy przedstawia rys. 29.5. Soczewka fluorytowa S (przepuszczająca promienie podczerwone), stanowiąca obiektyw lunetki L, skupia promieniowanie biegnące z wnętrza pieca na poczernionym spojeniu termoogniwa T. Przesłona D zapewnia padanie energii promienistej w określonym stożku o kącie bryłowym ω. Okular S_1 lunetki ustawiamy przez obrót śrubą regulacyjną R w takim położeniu, by otrzymać ostry obraz pieca. Wtedy promieniowanie pieca zostaje skupione na spojeniu termoogniwa i powoduje — zależne od temperatury pieca — wychylenie galwanometru włączonego do obwodu termoogniwa, nie zaznaczonego na rysunku. Galwanometr jest poprzednio wycechowany przy użyciu ciała doskonale czarnego o regulowanej i dokładnie znanej w każdej chwili temperaturze, tak że bezpośrednio na podziałce miernika elektrycznego odczytuje się wysokość temperatury.

Pirometry optyczne różnych typów pozwalają na pomiar temperatur do około 2500°C. W przypadku najwyższych mierzonych w ten sposób temperatur stosuje się osłabiacze, zmniejszające w określonym stopniu natężenie promieniowania.

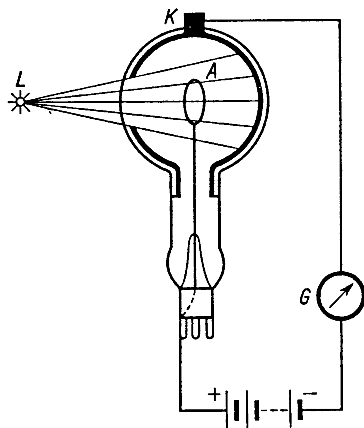
Opisana metoda zastosowana do ciała nieczarnego nie daje dokładnych wyników: Stosując np. pirometr Féry'ego wyznacza się nie rzeczywistą temperaturę ciała badanego, lecz tzw. *temperaturę promieniowania*. Jest to temperatura, jaką miałoby ciało czarne emitujące promieniowanie o ogólnej mocy równej całkowitej mocy wysyłanej przez rozważane ciało.

29.7. Zjawisko fotoelektryczne

Mniej więcej w tym samym czasie (około r. 1900), gdy starano się usunąć „katastrofę nadfioletową", odkryte zostało zjawisko, nazwane fotoelektrycznym, którego wyjaśnienie na gruncie fizyki klasycznej również stwarzało trudności nie do pokonania. Rozpoczniemy od charakterystyki tego zjawiska od strony doświadczalnej.



Rys. 29.6



Rys. 29.7

Zjawisko fotoelektryczne polega na wysyłaniu elektronów z powierzchni metali oświetlanych odpowiednim rodzajem promieniowania. Można to wykazać w następujący sposób. Płytkę cynkową Zn (rys. 29.6) o świeżo oczyszczonej powierzchni łączymy z elektroskopem i elektryzujemy ujemnie. Na płytkę kierujemy wiązkę promieni bogatą w nadfiolet, a więc np. z lampy łukowej *L* lub rtęciowej. Stwierdzamy rozładowywanie się elektroskopu, co jest równoznaczne ze zmniejszaniem się ładunku ujemnego na płytce. Mówimy, że pod wpływem światła z ujemnie naładowanej płytki „uchodzą" tzw. fotoelektrony. Jest to tzw. zjawisko fotoelektryczne.

Z płytki cynkowej silnie naładowanej dodatnio fotoelektrony nie są wysyłane, gdyż przeciwstawia się temu przyciąganie elektryczne ładunków różnoimiennych.

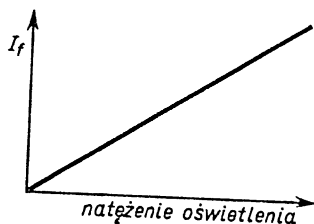
Zmiana źródła światła na źródło światła widzialnego (bez UV) lub zastosowanie przesłony absorbującej nadfiolet powoduje zanik zjawiska fotoelektrycznego na po-

wierzchni cynku, co wskazuje wyraźnie na wpływ rodzaju światła wywołującego zjawisko fotoelektryczne w danym materiale.

Zmieniając rodzaj metalicznej płytki stwierdzono, że np. w przypadku metali alkalicznych, takich jak sód, potas, rubid, cez do wywołania zjawiska wystarcza promieniowanie widzialne.

Do systematycznych badań doświadczalnych można zastosować obwód taki jak na rys. 29.7. Podstawową część obwodu stanowi komórka fotoelektryczna próżniowa K składająca się z bańki szklanej, z której dokładnie wypompowano powietrze. Wewnętrzna powierzchnia bańki, z wyjątkiem niewielkiego pola, pokryta jest cienką monomolekularną warstwą metalu alkalicznego. Przez to „okienko"* wprowadza się do wnętrza bańki wiązkę promieni. Wewnątrz bańki znajduje się elektroda A w postaci pierścienia lub spirali. Łączymy ją z biegunem dodatnim baterii, stanowi więc anodę. Biegun ujemny baterii łączymy z metaliczną warstwą światłoczułą, która tym samym staje się katodą. Do obwodu włączamy dodatkowo miernik prądu o dużej czułości, np. galwanometr zwierciadlany. Dopóki na warstwę światłoczułą nie pada promieniowanie, galwanometr nie wykazuje przepływu prądu; próżnia w bańce jest dobrym izolatorem. Z chwilą naświetlenia katody rozpoczyna się fotoemisja, fotoelektrony emitowane przez katodę zbierane są na anodzie i powodują powstanie w obwodzie prądu fotoelektrycznego.

Dokładna analiza zjawiska fotoelektrycznego obejmuje zbadanie zależności natężenia prądu fotoelektrycznego (fotoprądu) od: 1) natężenia oświetlenia badanej powierzchni, 2) częstotliwości zastosowanego promieniowania oraz 3) napięcia między katodą i anodą.



Rys. 29.8

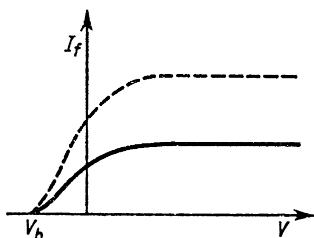
1. Wyniki badania *zależności natężenia prądu fotoelektrycznego I_f od natężenia oświetlenia badanej powierzchni* przy niezmiennej częstotliwości i niezmiennym napięciu między anodą A i katodą K przedstawione są na rys. 29.8. Liniowy przebieg krzywej wykresu świadczy o prostej proporcjonalności natężenia prądu do natężenia oświetlenia. Natężenie oświetlenia można zmieniać np. przez zmianę mocy danego źródła lub zmianę jego odległości (por. § 27.4).

2. Wyniki badania *zależności prądu fotoelektrycznego od częstotliwości promieniowania* (przy niezmienionych warunkach pozostałych) można ująć następująco: każdej

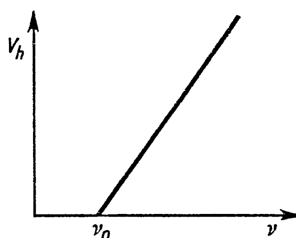
* W komórkach fotoelektrycznych przeznaczonych do pracy ze światłem nadfioletowym stosuje się okienka kwarcowe.

substancji emitującej elektrony pod wpływem światła można przypisać pewną *progową częstotliwość promieniowania* ν_0. Promieniowanie o częstotliwości mniejszej od progowej nie wywołuje efektu. Dla cynku częstotliwość progowa leży w nadfiolecie, dla Na, K, Rb, Cs — w dziedzinie promieniowania widzialnego. A więc częstotliwość progowa zależy od rodzaju naświetlanej substancji.

3. Wyniki badania *zależności prądu fotoelektrycznego od różnicy potencjałów V, przyłożonej między elektrody A i K*, przedstawione są na wykresie rys. 29.9. Ze wzrostem dodatniego potencjału przyspieszającego elektrony natężenie fotoprądu początkowo rośnie, a następnie osiąga wartość stałą, odpowiadającą prądowi nasycenia. W tych warunkach wszystkie fotoelektrony wyzwalane w jednostce czasu z powierzchni katody są w tym samym czasie zbierane na anodzie. Zmiana znaku różnicy potencjałów, czyli zastosowanie potencjału opóźniającego i wzrost jego wartości powodują zmniejszanie się prądu aż do jego zupełnego zaniku. Zanik prądu występuje dla danej częstotliwości promieniowania przy pewnej charakterystycznej dla danego materiału katody wartości potencjału, zwanej potencjałem hamowania V_h. Zaznaczona na tym samym wykresie krzywa przerywana (o większym prądzie nasycenia) dotyczy tego samego materiału katody (to samo V_h), lecz odpowiada silniejszemu oświetleniu powierzchni przy niezmiennej częstotliwości promieniowania.



Rys. 29.9



Rys. 29.10

Dodatkowe zbadanie zależności potencjału hamowania od częstotliwości promieniowania padającego stale na tę samą warstwę światłoczułą wykazuje proporcjonalność V_h do ν (rys. 29.10). Poniżej częstotliwości progowej ν_0 efektu fotoelektrycznego nie ma.

Dla odpowiedniej interpretacji zjawiska fotoelektrycznego bardzo ważne było doświadczenie, w którym radykalnie zmniejszano oświetlenie powierzchni katody przy $\nu > \nu_0$. Okazało się, że zjawisko fotoelektryczne, aczkolwiek bardzo słabe, występowało natychmiast po zastosowaniu oświetlenia i charakteryzowało się taką samą wartością V_h jak przy oświetleniach silnych.

Przejdźmy do wyjaśnienia podanych wyżej wyników doświadczalnych. Fakt, że nawet przy bardzo słabych oświetleniach zjawisko fotoelektryczne zachodzi natychmiastowo, jak również fakt istnienia charakterystycznej dla danego materiału częstotliwości progowej ν_0 przemawiają jednoznacznie za wyjaśnieniem tego zjawiska na podstawie teorii kwantowej promieniowania. Do wywołania zjawiska fotoelektrycznego potrzebny jest wkład energii na usunięcie elektronu z metalu (na wykonanie pracy wyjścia W_w — patrz § 32.5 i 32.6). Energii tej dostarcza kwant promieniowania $h\nu$.

Jego działanie jest skuteczne tylko wtedy, gdy $h\nu \geqslant W_w$, a więc istnieje pewna minimalna częstotliwość (progowa) ν_0 taka, że

$$h\nu_0 = W_w. \tag{29.5}$$

Jeśli wspomniana nierówność jest spełniona, to pochłonięty foton w natychmiastowym oddziaływaniu powoduje emisję elektronu. Każdy z fotonów działa we własnym zakresie; przy częstotliwości mniejszej od progowej energie poszczególnych kwantów nie dodają się i nie wytwarzają łącznej porcji energii zdolnej do wywołania fotoemisji.

Zobaczmy, jak stosując powyższe założenia można interpretować pozostałe wyniki doświadczalne. Proporcjonalność natężenia prądu fotoelektrycznego do natężenia oświetlenia jest prawie oczywista: im większe jest natężenie oświetlenia, tym więcej jest skutecznie działających kwantów promieniowania, a więc tym silniejszy jest prąd fotoelektryczny.

Omawiając w punkcie 3 wykres 29.9 wyjaśniliśmy już jego część odpowiadającą prądowi nasycenia. Wróćmy jednak do lewej części wykresu, a przede wszystkim do sensu fizycznego V_h. Z zasady zachowania energii wynika, że kosztem energii kwantu promieniowania $h\nu$ wykonana zostaje praca wyjścia fotoelektronu z metalu i — w przypadku nadwyżki $h\nu$ nad pracą wyjścia — nadana zostaje fotoelektronowi energia kinetyczna. *Potencjał hamowania* V_h stanowi potencjał potrzebny do zahamowania fotoelektronów o największej energii kinetycznej. Innymi słowy, praca związana z ruchem elektronu między punktami o różnicy potencjałów równej V_h (czyli eV_h) równa się $\frac{1}{2}mv^2_{max}$:

$$eV_h = \frac{1}{2}mv^2_{max},$$

gdzie v_{max} oznacza maksymalną prędkość początkową wyzwalanych fotoelektronów. Zastosowanie potencjału hamowania V_h między anodą i katodą nie dopuszcza fotoelektronów do anody. W miarę zmniejszania ujemnego potencjału anody coraz więcej fotoelektronów do niej dociera — stąd wzrost prądu fotoelektrycznego.

Podstawowe prawo rządzące zjawiskiem fotoelektrycznym zwane jest *prawem Einsteina*. Wyraża je następujące równanie:

$$h\nu = W_w + \frac{1}{2}mv_0^2. \tag{29.6}$$

Treść tego prawa można ująć następująco: *energia padającego kwantu promieniowania zostaje zużyta na wykonanie pracy wyjścia elektronu i na nadanie elektronowi energii kinetycznej.*

Jako *wnioski z tego prawa* można traktować: a) istnienie częstotliwości progowej ν_0, zależnej od rodzaju materiału fotokatody, b) zależność początkowej energii kinetycznej emitowanego fotoelektronu od częstotliwości padającego kwantu promieniowania, a niezależność jej od natężenia padającego promieniowania.

Korzystając z równania (29.5) można wzór (29.6) napisać w postaci

$$h\nu = h\nu_0 + \frac{1}{2}mv_0^2, \tag{29.7}$$

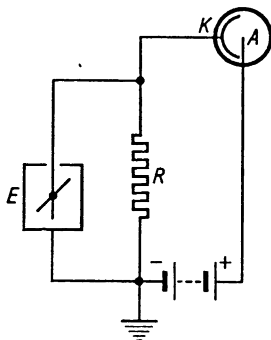
lub też, biorąc pod uwagę fotoelektrony o największej początkowej energii kinetycznej i opierając się na definicji potencjału hamującego, w postaci

$$eV_h = h\nu - h\nu_0 \quad \text{lub} \quad V_h = \frac{h}{e}(\nu - \nu_0). \tag{29.8}$$

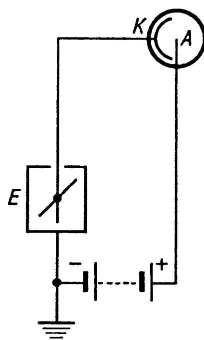
Na tej ostatniej liniowej zależności V_h od ν (rys. 29.10) można oprzeć metodę pomiaru stałej Plancka h. Tangens nachylenia prostej linii wykresu równa się bowiem ilorazowi h/e.

Szereg dokładnych pomiarów właściwości fotoelektrycznych różnych materiałów potwierdza całkowicie słuszność założeń Plancka i równania Einsteina.

Raz jeszcze podkreślamy, że wytłumaczenie zjawiska fotoelektrycznego możliwe jest tylko na gruncie teorii kwantowej. Każdy akt emisji fotoelektronu wywołany jest działaniem oddzielnego kwantu $h\nu$. W rozważaniach związanych ze zjawiskiem fotoelektrycznym jaskrawo zaznacza się nieciągłość struktury światła i niemożliwość wyjaśnienia zjawiska na gruncie teorii falowej.



Rys. 29.11



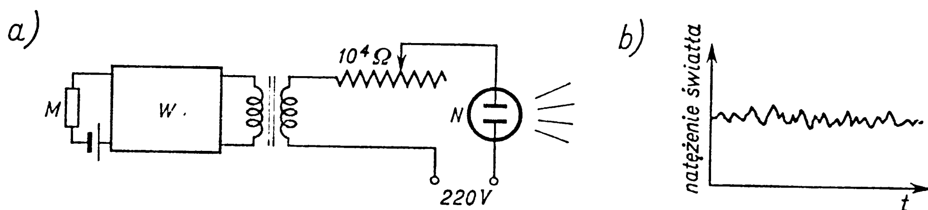
Rys. 29.12

Omówimy kilka zastosowań zjawiska fotoelektrycznego. Wśród zastosowań naukowych na pierwszy plan wysuwa się możliwość pomiaru małych natężeń światła. W obwodzie komórki fotoelektrycznej zastępujemy galwanometr elektrometrem E i mierzymy albo spadek napięcia na bardzo wielkim oporze R (rys. 29.11), albo czas ładowania się elektrometru (rys. 29.12). Stosując tę drugą metodę można mierzyć prądy fotoelektryczne o natężeniu 10^{-15} A, czyli około 10^4 razy słabsze niż mierzone bardzo czułymi galwanometrami. Czułość metody przekracza czułość oka ludzkiego. Znajduje ona zastosowanie w pomiarach astrofizycznych światła pochodzącego od gwiazd niedostrzegalnych gołym okiem.

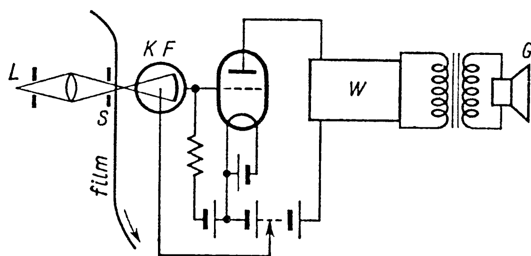
Ogólnie znane są zastosowania komórek fotoelektrycznych w urządzeniach liczących przedmioty nieprzezroczyste, które w swym ruchu przecinają poprzecznie strumień światła idący do fotokomórki. Można też z obwodem komórki połączyć urządzenia zabezpieczające lub sygnalizujące, np. na wypadek pożaru.

Opiszemy krótko jedną z metod nagrywania filmów dźwiękowych. Istota zagadnienia sprowadza się do wytworzenia światła, którego natężenie zmienia się analogicznie do zmian natężenia dźwięku i do utrwalenia tego „zmodulowanego" dźwiękiem światła na taśmie filmowej. Światło zmodulowane uzyskujemy przez doprowadzenie prądu powstającego w obwodzie mikrofonu M (po odpowiednim wzmocnieniu za pomocą wzmacniacza W) do obwodu neonówki N (rys. 29.13a). Gdy w obwodzie mikrofonowym

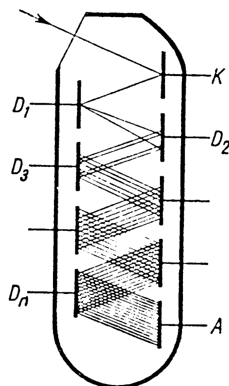
nie ma prądu, żarówka pali się stale równie jasno. Gdy fale głosowe dochodzą do mikrofonu, powstający w obwodzie mikrofonowym zmienny prąd nakłada się na stały prąd zasilający neonówkę i wywołuje zmiany w natężeniu wysyłanego przez nią promieniowania (rys. 29.13b). Podczas nagrywania filmu dźwiękowego takie zmodulowane światło neonówki kierowane jest przez bardzo wąską szczelinę na taśmę filmową przesuwającą się względem niej poprzecznie. Wtedy na bocznej części taśmy filmowej obok obrazków powstaje smuga o prążkach poprzecznych, wykazujących różny stopień zaczernienia. Jest to tzw. natężeniowa metoda nagrywania dźwięku na filmie.



Rys. 29.13



Rys. 29.14



Rys. 29.15

W czasie wyświetlania filmu bezpośrednio po przesunięciu taśmy przed obiektywem jest ona przesuwana ruchem jednostajnym przed wąską szczelinę *S* oświetloną zogniskowanym światłem żarówki *L* (rys. 29.14). Światło przechodzi przez różnie zaczernione prążki na taśmie filmowej i mniej lub więcej osłabione dochodzi do fotokomórki *KF*. Wahania natężenia światła wywołują odpowiednie wahania prądu w obwodzie fotokomórki. Prądy fotoelektryczne doprowadzane są do wzmacniacza *W*, a następnie w głośniku *G* przetwarzane są na dźwięki.

Na zjawisku fotoelektrycznym opiera się działanie przyrządu zwanego fotopowielaczem, zyskującego coraz szersze zastosowanie. Schemat fotopowielacza widzimy na rys. 29.15. Umieszczona w próżni światłoczuła katoda *K* wysyła fotoelektrony pod wpływem słabej wiązki padającego promieniowania. Elektrony te przyspieszane są odpowiednim napięciem, padają na powierzchnię wtórnej katody D_1, zwanej dynodą,

i wywołują wtórną emisję elektronową. Zwielokrotnienie strumienia elektronowego na skutek emisji wtórnej zależne jest od energii elektronów padających na dynodę oraz od rodzaju powierzchni dynody. Dla czystych powierzchni metalicznych *współczynnik zwielokrotnienia emisji* σ w korzystnych warunkach energetycznych dochodzi do 1,5. Zastosowanie powierzchni metalicznych pokrytych cienką warstwą półprzewodnika lub warstw złożonych, np. Cs_2O z domieszkami Ag i Cs, poważnie zwiększa wartość współczynnika emisji wtórnej, która w korzystnych warunkach energetycznych sięga do około 20. Zwielokrotniony na dynodzie D_1 strumień elektronów po odpowiednim przyspieszeniu pada na dynodę D_2, powtórnie ulega zwielokrotnieniu, pada na dynodę D_3 itd. Po n-krotnym wzmocnieniu na n-tej dynodzie pada na anodę A, z której wychodzi na zewnątrz w postaci prądu o natężeniu zwiększonym w stosunku do pierwotnego prądu fotoelektrycznego σ^n razy. Ostateczne wzmocnienie zależy w dużym stopniu od przyspieszającego napięcia między kolejnymi dynodami. W nowoczesnych fotopowielaczach w korzystnych dla nich warunkach pracy osiąga się wzmocnienia rzędu 10^6. Rozmiary nowoczesnych fotopowielaczy odpowiadają mniej więcej rozmiarom lamp elektronowych. Są one m.in. szeroko stosowane w fizyce jądrowej, jako urządzenie pomocnicze przy licznikach scyntylacyjnych.

29.8. Promienie Röntgena

29.8.1. *Ogólna charakterystyka właściwości promieni Röntgena.* Odkrycie promieni katodowych i badanie ich właściwości doprowadziło do nowego naukowego osiągnięcia. Okazało się, że po gwałtownym zahamowaniu promieni katodowych w ich ruchu od przeszkody, na której nastąpiło zahamowanie, rozchodzą się nowe promienie, zwane promieniami X lub promieniami Röntgena od nazwiska ich odkrywcy (rok 1895).

Już początkowe badania właściwości promieni Röntgena wykazały, że mają one zdolność:

1) wzbudzania fluorescencji licznych ciał (m.in. wywołują świecenie ekranów pokrytych platynocjankiem baru lub siarczkiem cynku),

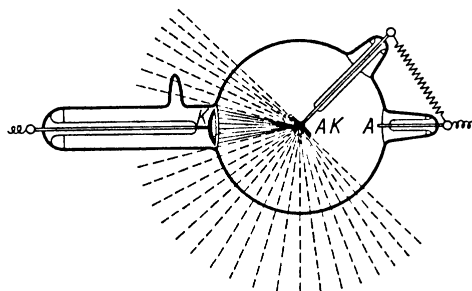
2) zaczerniania kliszy fotograficznej,

3) wywoływania jonizacji gazów (stwierdzono szybkie rozładowywanie się elektroskopów umieszczonych w pobliżu źródeł promieni Röntgena),

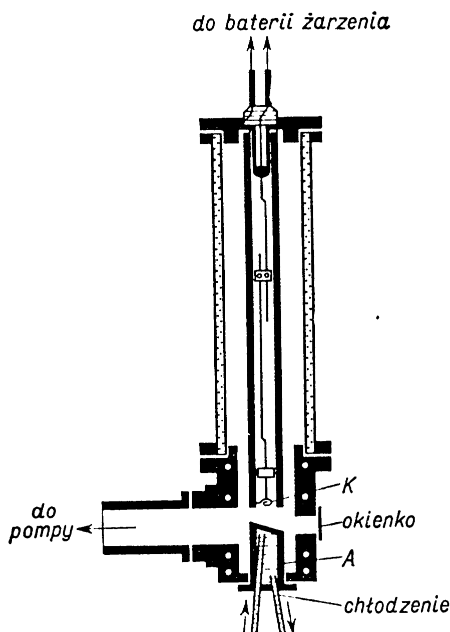
4) przenikania przez materię.

Przenikliwość promieni Röntgena zależy od warunków ich wytwarzania. Początkowo promienie Röntgena były wytwarzane w tzw. lampach rentgenowskich jonowych. Budowa takiej lampy przedstawiona jest schematycznie na rys. 29.16. W bańce próżniowej, w której ciśnienie waha się około 0,1 N/m^2 (0,001 mm Hg), znajduje się katoda K w postaci wklęsłej blaszki, np. z glinu, anoda A i antykatoda AK w postaci płytki z trudno topliwego metalu, np. wolframu lub platyny. Antykatoda umieszczona jest w odległości równej promieniowi krzywizny od katody. Dzięki temu wiązka elektronów wybiegająca z katody prostopadle do jej powierzchni skupia się na powierzchni antykatody. Antykatoda ma ten sam potencjał co anoda, więc w zasadzie oddzielna anoda nie jest potrzebna. Napięcie istniejące między anodą i katodą decyduje o prędkości

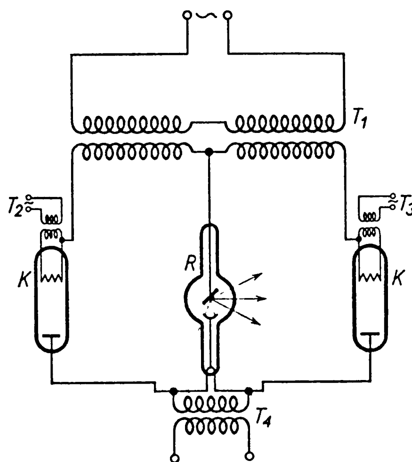
elektronów dobiegających do antykatody. Jest ono rzędu kilkudziesięciu tysięcy woltów. Wyższym napięciom odpowiadają *promienie Röntgena* bardziej przenikliwe — bardziej *twarde*. Niższym napięciom na elektrodach lampy rentgenowskiej odpowiadają promienie bardziej *miękkie*, łatwiej absorbowane przez materię.



Rys. 29.16



Rys. 29.17



Rys. 29.18

Ciśnienie panujące w lampie rentgenowskiej jonowej ma również wpływ na twardość promieni Röntgena. Z biegiem czasu ciśnienie w bańce się zmienia, gdyż część gazu zostaje zaadsorbowana na ściankach. W tych warunkach potrzebne jest wyższe napięcie do wzbudzenia promieni Röntgena i stają się one twardsze. Stałość ciśnienia można utrzymać przez wprowadzenie do bańki lampy rurki palladowej z okludowanym wodorem (*okluzja* — pochłanianie gazów przez niektóre ciała stałe), która po ogrzaniu uwalnia wodór i uzupełnia niedobór ciśnienia.

Rysunek 29.17 przedstawia nowoczesną lampę rentgenowską z żarzoną katodą, zwaną lampą Coolidge'a. Wewnątrz lampy mamy wysoką próżnię. Ciśnienie jest rzędu 10^{-7} N/m^2 (czyli 10^{-9} mm Hg). Katodę K stanowi spiralka wolframowa, anodą jest podstawa pustego miedzianego cylindra. Może to być płytka z platyny, wolframu, molibdenu, srebra. Miedziane ścianki cylindra zapewniają dobre odprowadzanie ciepła. W pewnych typach lamp stosuje się dodatkowe chłodzenie powietrzne, olejowe lub wodne. Temperatura żarzenia katody zależna jest od zastosowanego napięcia żarzenia. Im wyższa jest temperatura katody, tym większa jest gęstość prądu termoemisji i w dalszych skutkach — tym większe natężenie promieni Röntgena. Natężenie promieni Röntgena decyduje o intensywności świecenia ekranów fluoryzujących, o stopniu zaczernienia kliszy fotograficznej, o stopniu jonizacji gazów, co jest stosunkowo najłatwiejsze do mierzenia. Napięcie włączone między anodę i katodę decyduje o twardości promieni Röntgena. Źródłem tego napięcia jest transformator wysokiego napięcia T_1. Przed doprowadzeniem napięcia do zacisków lampy R zostaje ono wyprostowane przez lampę prostowniczą K (ewentualnie dwie lampy), zwaną kenotronem. Rysunek 29.18 przedstawia schemat połączeń nowoczesnej lampy rentgenowskiej. Transformatory T_2, T_3 i T_4 są źródłami napięcia żarzenia w lampach prostowniczych i lampie rentgenowskiej (lampa rentgenowska zaznaczana jest na schematach jako lampa jonowa).

Przenikliwość promieni Röntgena zależy nie tylko od napięcia, przy którym pracuje lampa, lecz także od rodzaju materii, przez którą przechodzą promienie. Okazało się, że czynnikiem decydującym o absorpcji promieni Röntgena jest numer porządkowy Z (zwany też liczbą atomową) pierwiastka w układzie Mendelejewa i grubość warstwy ciała absorbującego. Jeżeli przez I_0 oznaczymy natężenie wiązki padającej, a przez I — natężenie wiązki wychodzącej z warstwy absorbującej o grubości x, to

$$I = I_0 e^{-\mu x}, \tag{29.9}$$

gdzie μ oznacza *liniowy współczynnik absorpcji* danego środowiska. Jest on funkcją numeru porządkowego pierwiastka absorbującego i wyraża logarytm naturalny ułamka, o jaki zmniejsza się natężenie promieni Röntgena po przejściu jednostkowej drogi w ośrodku absorbującym. (Pamiętać jednak należy, że μ jest ponadto zależne od rodzaju padających promieni Röntgena.)

Wzór (29.9) otrzymujemy w wyniku prostych rozważań matematycznych. Zmiana natężenia wiązki po przejściu przez warstwę absorbującą jest proporcjonalna do natężenia I oraz do grubości warstwy dx, czyli

$$-dI = \mu I dx,$$

$$-\frac{dI}{I} = \mu dx,$$

a po scałkowaniu

$$\ln I = -\mu x + \text{const.}$$

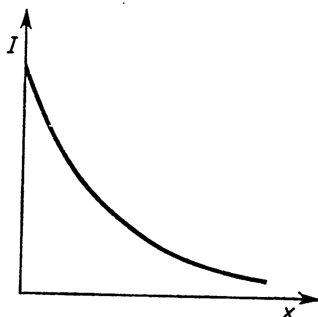
Wartość stałej całkowania znajdujemy z warunku, że dla $x = 0$ $I = I_0$, a zatem

$$\ln I = -\mu x + \ln I_0,$$

$$\ln\frac{I}{I_0} = -\mu x, \quad \text{czyli} \quad I = I_0 e^{-\mu x}.$$

Wykres tej zależności przedstawiony jest na rys. 29.19.

Zależność absorpcji promieni Röntgena od rodzaju ciała absorbującego umożliwia stosowanie ich do celów diagnostycznych w medycynie. Ciało ludzkie absorbuje słabiej niż kości, a te ostatnie słabiej niż np. odłamki pocisków lub kule tkwiące w organizmie. Wprowadzenie silnie absorbujących „papek" do przewodu pokarmowego lub tzw. płynów kontrastowych do płuc umożliwia otrzymywanie zdjęć rentgenowskich (rentgenogramów), stanowiących podstawę do stawiania diagnozy lub kontroli stanu zdrowia. Na tej samej podstawie opiera się tzw. defektoskopia rentgenowska: prześwietlanie bloków metalicznych pozwala na kontrolę np. odlewów, wykrywa wady w postaci pęcherzyków powietrza, pęknięć wewnętrznych itp.



Rys. 29.19

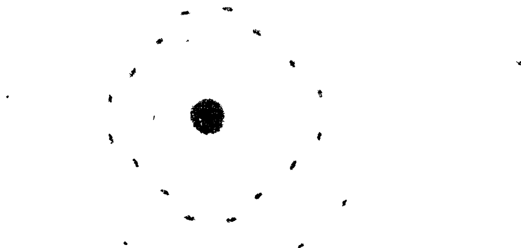
Lampy rentgenowskie stosowane do celów diagnostycznych w medycynie pracują pod napięciem około 60 000 V i przy natężeniu prądu 5–400 mA. Przy większym natężeniu prądu stosuje się krótkie czasy pracy lampy, około 5–10 sekund. W technice stosowane są lampy niskonapięciowe (30 000–80 000 V) i wysokonapięciowe do 250 000 V. Iloczyn UI przedstawia moc lampy. Na przykład dla lampy o napięciu 100 000 V i natężeniu prądu 5 mA moc wynosi 500 W. Z tego bardzo mały ułamek, bo zaledwie około $\frac{1}{1000}$ przekształca się na energię promieni Röntgena, reszta zamienia się na ciepło. Liczby te wskazują na konieczność stosowania trudno topliwych metali jako materiału na antykatody i intensywnego chłodzenia zależnego ponadto od tego, czy lampa jest przewidziana do pracy ciągłej, czy w krótkich impulsach.

W czasie pracy z lampami rentgenowskimi konieczne jest zachowanie wielu środków ostrożności, gdyż absorpcja promieni Röntgena przez ciało ludzkie powoduje zmiany we krwi, oparzenia rentgenowskie itp. Działanie jest tym niebezpieczniejsze, że efekty krótkich, powtarzających się naświetlań kumulują się w organizmie prowadząc do poważnych zmian. Z tego powodu lampy umieszcza się w osłonach ołowianych, promienie wychodzą na zewnątrz tylko przez okienka o możliwie najmniejszych rozmiarach, a więc jako wiązki o określonym kierunku przebiegu. Ekrany do prześwietlań rentgenowskich mają od strony obserwatora zabezpieczenie w postaci grubych warstw ze szkła ołowiowego. Dla zabezpieczenia przed atakującym ze wszystkich stron pro-

mieniowaniem rentgenowskim rozproszonym na cząstkach powietrza, częściach aparatury itp. pracownicy noszą specjalne fartuchy ochronne gumowe z dodatkiem ołowiu.

Promienie Röntgena działają niszcząco na młode tkanki. Zostało to wykorzystane przy zwalczaniu szybko rozwijającej się tkanki nowotworów.

29.8.2. *Natura falowa promieni Röntgena.* Dalsze badania promieni Röntgena przyniosły wiadomości o ich charakterze. Doświadczenia wykazały, że promienie Röntgena nie ulegają odchyleniu od pierwotnego biegu ani w polu elektrycznym, ani w polu magnetycznym. Nie są zatem cząsteczkami obdarzonymi ładunkami elektrycznymi, nie są odbitymi od powierzchni antykatody elektronami. Nasunęło się przypuszczenie co do ich charakteru falowego. Potwierdzeniem tej hipotezy byłoby doświadczalne stwierdzenie w dziedzinie promieniowania rentgenowskiego charakterystycznych zjawisk falowych, takich jak ugięcie, interferencja, polaryzacja. Dopiero w r. 1912 Laue, a w r. 1913 Braggowie uzyskali efekty interferencji promieni Röntgena po zastosowaniu całkowicie różnych metod. Trudności eksperymentalne wiązały się z tym, że długość fali promieni Röntgena jest tak mała, że zawodziły zwykłe metody badań optycznych. Do wywołania ugięcia nie nadawały się zwykłe siatki dyfrakcyjne ustawiane prostopadle do kierunku biegu promieni. Potrzebne były takie siatki, których stałe byłyby rzędu odległości międzyatomowych w ciałach stałych.

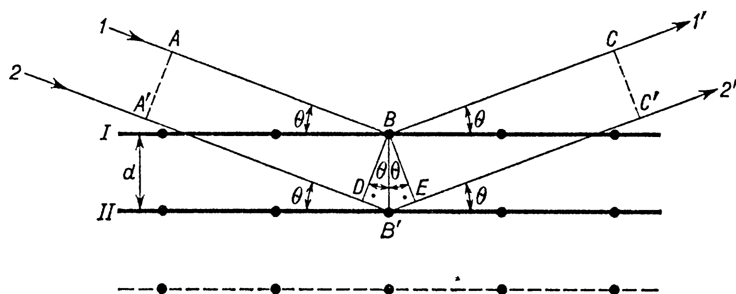


Rys. 29.20

Laue pierwszy wysunął myśl, że do ugięcia i interferencji promieni Röntgena można wykorzystać regularny układ elementów kryształu, czyli atomów, jonów lub cząsteczek (por. § 16.12). Doświadczenia tego typu wykonali natępnie Friedrich i Knipping. Wąska wiązka promieni Röntgena przechodzi przez cienką płytkę wyciętą z kryształu, a następnie pada na kliszę fotograficzną. Na kliszy otrzymujemy zaczernienie w miejscu centralnym odpowiadającym nieugiętemu przebiegowi promieni oraz szereg zaczernień (plamek) rozmieszczonych z zachowaniem pewnej symetrii dokoła plamki centralnej (rys. 29.20). Obraz interferencyjny zmienia swą symetrię przy zmianie rodzaju

płytki krystalicznej. Takie rentgenogramy były już dowodem ugięcia i interferencji promieni Röntgena, lecz były trudne do bliższej analizy w związku z trójwymiarowością użytych siatek krystalicznych. W każdym razie były to już dowody falowej natury promieni Röntgena, jak również dowody występowania regularności w budowie kryształów.

Interesującą metodę interferencyjną badania zarówno długości fali promieni Röntgena, jak i struktury kryształów opracowali *Braggowie*. Rozpatrywali oni wyniki interferencji promieni Röntgena *rozproszonych* na elementach kryształu znajdujących się na jego powierzchni i na równoległych do niej dalszych płaszczyznach najgęstszego rozmieszczenia tych elementów. Będziemy je nazywać *płaszczyznami strukturalnymi kryształu*. Pokrywają się one z płaszczyznami łupliwości kryształu. Rysunek 29.21 przedstawia m.in. takie dwie sąsiednie płaszczyzny *I* i *II* w krysztale. Kropki oznaczają położenia elementów kryształu. Każdy z elementów rozprasza promienie rentgenowskie na wszystkie strony. Promienie rozproszone nakładają się na siebie i na ogół wygaszają się wzajemnie. Tylko w przypadku spełnienia pewnych warunków pojawia się wzmocnienie interferujących promieni, a mianowicie wzmacniają się tylko te promienie rozproszone, które: 1) podlegają prawom odbicia, tzn. leżą w płaszczyźnie padania i tworzą z normalną do powierzchni rozpraszającej kąt równy kątowi padania, 2) powstają wskutek rozproszenia na różnych płaszczyznach strukturalnych kryształu (np. *I* i *II* na rys. 29.21) i 3) różnice ich dróg optycznych przy nakładaniu się równe są całkowitej wielokrotności fali.



Rys. 29.21

Treść warunku 1 może prowadzić do pewnych nieporozumień. Trzeba sobie dobrze uświadomić, że aczkolwiek warunek ten pokrywa się z prawami odbicia w optyce geometrycznej, to jednak w metodzie Braggów istotną rolę odgrywa rozpraszanie, a nie odbicie promieni Röntgena na elementach kryształu. Dlatego też, jeśli chcemy mimo wszystko nawiązywać do zjawiska odbicia, to metodę Braggów należy nazywać metodą *selektywnego odbicia*, celem podkreślenia specyficznych cech rozproszonych promieni wzmacniających się podczas interferencji.

W celu wyprowadzenia matematycznego warunku wzmocnienia interferujących promieni rozpatrzmy promienie AB i $A'B'$, ulegające rozproszeniu na elementach krystalicznych płaszczyzn strukturalnych *I* i *II* (rys. 29.21). Z warunku 1 wynika,

że ze wszystkich możliwych rozproszonych w punktach B i B' promieni wzmocnieniu będą mogły ulec tylko promienie BC i $B'C'$, spełniające prawa odbicia. Z warunku 3 natomiast wynika powiązanie różnicy dróg optycznych obu tych promieni (równej $DB'+B'E$) z długością fali λ, a mianowicie

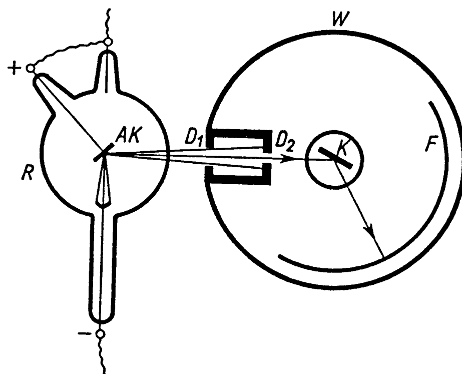
$$n\lambda = DB'+B'E,$$

czyli

$$n\lambda = 2d\sin\theta, \quad \text{przy czym} \quad n = 1, 2, 3, \ldots \tag{29.10}$$

gdzie d oznacza odległość sąsiednich płaszczyzn strukturalnych, będącą równocześnie *stałą siatki krystalicznej* (wielkość rzędu 10^{-10} m), θ — tzw. *kąt odblysku*, stanowiący dopełnienie kąta padania do 90°, n — rząd widma.

Fale rozproszone na wszystkich równoległych płaszczyznach atomowych nie będą się znosiły podczas nakładania się, lecz przeciwnie będą się wzmacniały, gdy będzie spełniony warunek wyrażony równaniem (29.10), zwany *warunkiem Bragga*. Warunek Bragga jest warunkiem koniecznym do obserwowania wzmocnienia interferencyjnego rozproszonych promieni Röntgena. Jasne jest bowiem, że gdy różnica dróg równa się $(2n+1)\frac{\lambda}{2}$, fale pochodzące od każdej pary sąsiednich płaszczyzn wygaszają się przy interferencji. Jeżeli różnica dróg nie będzie równa ani parzystej, ani nieparzystej wielokrotności $\lambda/2$, to wygaszeniu ulegną nie fale pochodzące od 2 sąsiednich płaszczyzn, lecz od 1 i jakiejś m-tej, od 2 i $(m+1)$-ej itd., tzn. od takich płaszczyzn, że całkowita różnica dróg będzie równa nieparzystej wielokrotności $\lambda/2$.

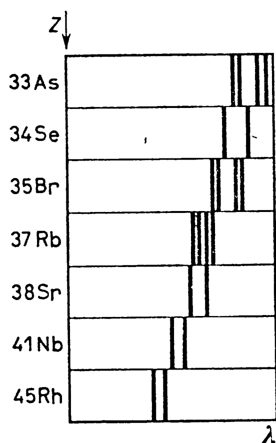


Rys. 29.22

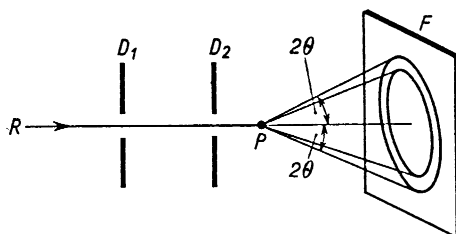
Znając stałą siatki danego kryształu i kąt odbłysku można wyznaczać długości fal rentgenowskich. Otrzymuje się wielkości rzędu 10^{-8}–10^{-11} m.

Do wyznaczania długości fal rentgenowskich służą spektrografy rentgenowskie. Rysunek 29.22 przedstawia schematycznie taki spektrograf (widok z góry). Widzimy lampę R, przesłony D_1 i D_2 wycięte z ołowiu przepuszczające wąską wiązkę promieni Röntgena, płytkę krystaliczną K, umieszczoną na ruchomym stoliku spektrografu. Stolik z płytką znajduje się w środku walcowatej puszki W. Na wewnętrznej części

powierzchni bocznej puszki umieszczona jest błona fotograficzna. Stolik z kryształem zostaje wprawiony w powolny ruch obrotowy za pomocą mechanizmu zegarowego. Załóżmy, że w wiązce padającej na kryształ występują fale o kilku określonych długościach fali. Każdej z długości fal odpowiada określony kąt odbłysku. Przy pewnym położeniu kryształu spełniony jest warunek Bragga, np. dla fali o długości λ_1. Fale te wzmacniają się wskutek interferencji i dają na błonie fotograficznej (po wywołaniu) prążek odpowiadający λ_1, analogiczny do prążków w widmie światła widzialnego. Przy innym ustawieniu płytki (dalszy obrót) spełniony będzie warunek Bragga dla λ_2 i w innym miejscu błony powstanie prążek reprezentujący falę λ_2. Ostatecznie po wywołaniu błony widoczne są na niej ostro zarysowane prążki na słabym ciągłym tle. Zmieniając antykatodę otrzymujemy inny układ prążków, gdyż inne długości fal rentgenowskich są charakterystyczne dla nowego materiału antykatody. Zmieniając kryształ wprowadzamy inną stałą siatki *d*, a tym samym zmieniamy kąty odbłysku dla danych długości fal. Przykład rozmieszczenia linii rentgenowskich dla kilku pierwiastków pokazany jest na rys. 29.23. Na rysunku tym nie jest zaznaczone tło ciągłe, a tylko podane są położenia charakterystycznych linii rentgenowskich, tzw. serii *K* (por. § 30.12).



Rys. 29.23



Rys. 29.24

W późniejszej metodzie badania widm promieni Röntgena, opracowanej przez *Debye'a* i *Scherrera*, stosowano zamiast pojedynczego obracanego kryształu (monokryształu) pręcik ze sprasowanego proszku krystalicznego, a więc substancji, o której można zakładać, że występuje w niej całkowicie chaotyczny układ bardzo dużej liczby mikrokryształów. Dzięki temu bezpośrednio, bez dodatkowego obrotu stolika, otrzymywano wszystkie możliwe ustawienia płaszczyzn strukturalnych w krysztale. Zdjęcia robiono na filmie cylindrycznym lub płaskim. Często też stosowano zamiast pręcików mikrokrystaliczne folie metaliczne, tzn. cieniutkie warstwy metaliczne, złożone z bezładnie rozmieszczonych mikroskopijnych ziaren krystalicznych. Rysunek 29.24 przedstawia schematycznie układ doświadczalny w metodzie Debye'a i Scherrera. Z nie zaznaczonej na rysunku lampy *R* promienie Röntgena po przejściu przez przesłony

D_1 i D_2, wycięte w płytkach ołowianych, padają w postaci wąskiej wiązki na pręcik P, a następnie na film płaski F ustawiony prostopadle do pierwotnego kierunku biegu wiązki. Wobec najzupełniej chaotycznego rozkładu strukturalnych płaszczyzn mikrokryształów w pręciku wszystkie promienie rozproszone spełniające warunek Braggów zajmują położenia w przestrzeni odpowiadające powierzchni bocznej stożka o kącie przy wierzchołku 4θ i osi leżącej na przedłużeniu kierunku wiązki padającej. Wynika to stąd, że każdy promień rozproszony pod kątem odbłysku θ tworzy z kierunkiem przedłużenia promienia padającego kąt 2θ (por. rys. 29.21). Padając na błonę fotograficzną promienie te wywołują zaczernienia w postaci koncentrycznych pierścieni. Jednostajne zaczernienie wzdłuż całego obwodu każdego z pierścieni świadczy o idealnie chaotycznym rozkładzie mikrokryształów. Istnienie uprzywilejowanych kierunków uporządkowania mikrokryształów zaznacza się na rentgenogramach w postaci szczególnie silnych zaczernień w pewnych miejscach filmu, a nieciągłości pierścieni — w innych.

Otrzymane w opisany sposób rentgenogramy można wykorzystać do badania:

1) długości fali promieni Röntgena (stosując kryształy o znanej strukturze, tzn. o znanych odległościach d),

2) struktury kryształów (stosując promienie Röntgena o znanej długości fali λ),

3) uporządkowania mikrokryształów w drutach, foliach itp. W badaniach krystalograficznych wygodnie jest posługiwać się monochromatycznym promieniowaniem rentgenowskim. Wiązkę monochromatyczną można otrzymać stosując jako wstępny etap badań rozproszenie metodą Braggów pod kątem θ. Można też do tego celu stosować odpowiednie filtry rentgenowskie.

Późniejsze badania wykazały, że do analizy promieni Röntgena można też wykorzystywać najprecyzyjniejsze siatki dyfrakcyjne optyczne stosując prawie styczne padanie wiązek na powierzchnię siatki.

Bliższą analizę widm rentgenowskich i wyjaśnienie mechanizmu ich powstawania podamy w § 30.12.

Pytania i zadania

1. Jakie ciało nazywamy doskonale czarnym? Podać przykłady ciał o właściwościach zbliżonych do ciała doskonale czarnego.

2. Jakie prawa rządzą promieniowaniem termicznym ciała doskonale czarnego?

3. W jakim prawie występuje zdolność emisji i zdolność absorpcji danego ciała? Podać treść tego prawa oraz definicje wymienionych wielkości.

4. Na jakim tle powstały trudności teorii elektromagnetycznej światła? Jaką nową koncepcję mechanizmu emisji promieniowania wprowadził Planck? Jakie fakty doświadczalne można wyjaśnić przyjmując słuszność tej koncepcji?

5. Na czym polega zjawisko fotoelektryczne?

6. W jaki sposób można zmieniać częstotliwość progową, potencjał hamowania, wartość prądu nasycenia?

7. Omówić kilka zastosowań zjawiska fotoelektrycznego.

8. W jaki sposób można wykorzystać zjawisko fotoelektryczne do wyznaczania stałej Plancka?

9. Czym się różni sposób wytwarzania promieni Röntgena za pomocą: a) lamp jonowych, b) lamp Coolidge'a?

10. Jakie są podstawowe właściwości promieni Röntgena?

11. Jakie są dowody natury falowej promieni Röntgena?

12. Czy słuszne jest twierdzenie, że metoda Braggów opiera się na zjawisku odbicia światła? Dokładnie uzasadnić odpowiedź.

13. Czym się różni metoda obracanego monokryształu od metody Debey'a? Jakie wiadomości można otrzymać z rentgenogramów otrzymywanych metodą Debye'a?

14. Włókno wolframowe o promieniu 0,1 mm i długości 20 cm, znajdujące się w próżni, ogrzano do temperatury 3000K. Zakładając, że włókno zachowuje się jak ciało doskonale czarne obliczyć ogólną ilość energii, którą ono wypromieniowuje w ciągu 30 s?

Odp. 17 400 J.

15. Zwykle przyjmuje się, że 1 cm^2 powierzchni Ziemi traci w ciągu 1 min wskutek promieniowania średnio 0,5447 J. Jaką temperaturę musiałoby mieć ciało doskonale czarne wypromieniowujące tę samą ilość energii?

Odp. −73°C.

16. Zakładając, że Słońce promieniuje tak, jak ciało doskonale czarne, obliczyć temperaturę jego powierzchni przy założeniu, że maksimum natężenia wysyłanego przez nie promieniowania przypada na długość fali 5000 Å.

Odp. 5780K.

17. Włókno wolframowe o średnicy $d = 0{,}2$ mm ogrzewa się wskutek przepływu prądu elektrycznego do temperatury 3000K. W ciągu jakiego czasu włókno oziębi się do temperatury 1000K, jeżeli założymy, że nie pobiera ono wcale ciepła od otoczenia, a traci wyłącznie przez promieniowanie, tak jak ciało doskonale czarne. Gęstość wolframu 19 g/cm^3, ciepło właściwe 155 J/kg · K.

Odp. 0,828 sekund.

18. Obliczyć energię 1 fotonu dla światła czerwonego o długości fali 8000 Å i dla promieniowania rentgenowskiego o długości fali 1 Å. Obliczyć stosunek obu tych energii.

Odp. $2{,}48 \cdot 10^{-19}$ J, $19{,}86 \cdot 10^{-16}$ J, 1:8000.

19. Żarówka lampki kieszonkowej pobiera moc około 1 W. Zakładając, że promieniowanie rozchodzi się bez absorpcji równomiernie we wszystkich kierunkach w przestrzeni i że średnia długość wysyłanej fali równa się 1 μm, obliczyć liczbę fotonów padających w ciągu 1 s na 1 cm^2 powierzchni ustawionej prostopadle wzgiędem biegu promieni w odległości 1 km.

Odp. $4 \cdot 10^7$ fotonów/cm^2 · s.

20. Obliczyć graniczną długość fali światła, które może wywołać zjawisko fotoelektryczne w platynie i cezie. Praca wyjścia elektronu z platyny 6,3 eV, z cezu 1,9 eV (1 eV = $1{,}602 \cdot 10^{-19}$ J).

Odp. $1{,}97 \cdot 10^{-7}$ m; $6{,}53 \cdot 10^{-7}$ m.

21. Z jaką prędkością początkową wylatują fotoelektrony z płytki srebrnej, oświetlonej światłem monochromatycznym o długości fali $2 \cdot 10^{-7}$ m, jeśli długość fali, odpowiadająca częstości progowej, dla srebra wynosi $2{,}6 \cdot 10^{-7}$ m. Masa elektronu $9 \cdot 10^{-31}$ kg.

Odp. $7{,}09 \cdot 10^5$ m/s.

22. Na kryształ chlorku potasu padają promienie rentgenowskie i dają widmo drugiego rzędu pod kątem 18°03′. Pod jakim kątem wytworzy się widmo pierwszego rzędu?

Odp. 8°55′.

23. Jaka jest stała siatki chlorku potasu, jeżeli w warunkach takich jak w zad. 22 zastosowano promienie rentgenowskie o długości fali 1,537 Å?

Odp. $4{,}96 \cdot 10^{-10}$ m.

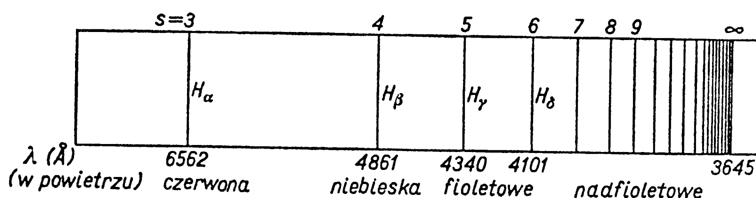
ROZDZIAŁ 30

Fizyka atomowa

30.1. Widmo atomu wodoru

W punkcie 28.5.1 wspominaliśmy o atomowym pochodzeniu widm liniowych, charakterystycznych dla poszczególnych pierwiastków. W paragrafie niniejszym rozpatrzymy dokładniej zależność widm liniowych od struktury atomu.

Wśród widm liniowych najprostszy charakter ma widmo wodoru. Zastosujmy jako źródło promieniowania rurkę Geisslera z rozrzedzonym wodorem. Załączając elektrody rurki do iskierników cewki indukcyjnej Ruhmkorffa, wzbudzamy promieniowanie. Po ustawieniu rurki przed soczewką kolimatora spektroskopu otrzymujemy w polu widzenia lunety kilka charakterystycznych dla wodoru linii w dziedzinie promieniowania widzialnego. Linie te oznaczono symbolami H_α, H_β, H_γ itd. i zmierzono ich długości fal (rys. 30.1).



Rys. 30.1

Linie te były m.in. przedmiotem zainteresowania Balmera, który usiłował znaleźć wzór matematyczny wyrażający długości ich fal. Ostatecznie stwierdził on, że obowiązuje zależność następująca:

$$\lambda = 3645{,}6\frac{n^2}{n^2-4}\ \text{Å}.$$

W latach późniejszych nadano temu wzorowi inną postać, a mianowicie

$$\frac{1}{\lambda} = R_{\mathrm{H}}\left(\frac{1}{2^2}-\frac{1}{s^2}\right), \qquad (30.1)$$

gdzie $s = 3, 4, 5, \ldots$

Wielkość R_H nosi nazwę *stałej Rydberga* (dla wodoru) i ma wartość 10 967 758 m^{-1}, gdy długość fali wyrażona jest w metrach; $1/\lambda$ jest to tzw. *liczba falowa* $\tilde{\nu}$, która określa, ile długości fal mieści się na odcinku 1 m.

Omawiane linie tworzą tzw. *serię Balmera*. Jak widać z rys. 30.1, długości fal tej serii stopniowo coraz mniej różnią się od siebie. Linie tworzą ciąg zgęszczający się w otoczeniu linii granicznej. Długość fali linii granicznej znajdujemy, podstawiając do wzoru Balmera zamiast s nieskończoność. Graniczna długość fali serii Balmera wynosi około 3650 Å.

Późniejsze badania widma wodoru przyniosły wiadomości o innych seriach znajdujących się już w dziedzinie promieniowania niewidzialnego. Nadano im nazwy zgodne z nazwiskami ich odkrywców. Tak więc seria wodorowa w nadfiolecie nosi nazwę *serii Lymana*. Seria Balmera obejmuje dziedzinę widzialną i bliski nadfiolet, serie *Paschena*, *Bracketta*, *Pfunda* i *Humphreysa* (ta ostatnia wykryta w r. 1952) leżą w podczerwieni. Okazało się, że poszczególnym seriom można przyporządkować wzory podobne do wzoru Balmera, a mianowicie:

seria Lymana (1916)	$\frac{1}{\lambda} = R_H\left(\frac{1}{1^2} - \frac{1}{s^2}\right)$,	$s = 2, 3, 4, \ldots$
seria Balmera (1885)	$\frac{1}{\lambda} = R_H\left(\frac{1}{2^2} - \frac{1}{s^2}\right)$,	$s = 3, 4, 5, \ldots$
seria Paschena (1908)	$\frac{1}{\lambda} = R_H\left(\frac{1}{3^2} - \frac{1}{s^2}\right)$,	$s = 4, 5, 6, \ldots$
seria Bracketta (1922)	$\frac{1}{\lambda} = R_H\left(\frac{1}{4^2} - \frac{1}{s^2}\right)$,	$s = 5, 6, 7, \ldots$
seria Pfunda 1924)	$\frac{1}{\lambda} = R_H\left(\frac{1}{5^2} - \frac{1}{s^2}\right)$,	$s = 6, 7, \ldots$
seria Humphreysa (1952)	$\frac{1}{\lambda} = R_H\left(\frac{1}{6^2} - \frac{1}{s^2}\right)$,	$s = 7, 8, \ldots$

Wobec analogicznej postaci wzorów wszystkich serii można napisać ogólny wzór dla serii wodorowych

$$\frac{1}{\lambda} = R_H\left(\frac{1}{n^2} - \frac{1}{s^2}\right), \tag{30.2}$$

gdzie $n = 1, 2, 3, 4 \ldots$ jest charakterystyczne (stałe) dla danej serii, $s = n+1, n+2, n+3, \ldots$ określa poszczególne linie danej serii. Często wzór (30.2) podawany jest w innej postaci, a mianowicie

$$\tilde{\nu} = \frac{R_H}{n^2} - \frac{R_H}{s^2} = T_n - T_s, \tag{30.3}$$

gdzie $\tilde{\nu}$ oznacza liczbę falową, a $T_n = \frac{R_H}{n^2}$ i $T_s = \frac{R_H}{s^2}$ tzw. *termy*. Inaczej mówiąc,

liczbę falową określonej linii widmowej danej serii wyznacza różnica dwóch termów, z których odjemna jest *termem stałym*, charakterystycznym dla danej serii, a odjemnik jest termem zmiennym, zwanym często *termem bieżącym*. Łatwo sprawdzić, że term stały odpowiada liczbie falowej granicy krótkofalowej danej serii $\left(\text{gdy}\quad s \to \infty,\ \text{to}\ \frac{1}{\lambda} \to \frac{R_H}{n^2}\right)$.

Wzory wiążące linie wodoru w serie są bardzo proste. Nasuwa się pytanie, jaki mechanizm emisji poszczególnych linii widmowych prowadzi do tak prostych zależności. Odpowiedź na to pytanie wymaga pogłębienia wiadomości dotyczących budowy atomu. Rozpoczniemy od podkreślenia pewnych etapów rozwoju tej dziedziny fizyki.

30.2. Modele atomu

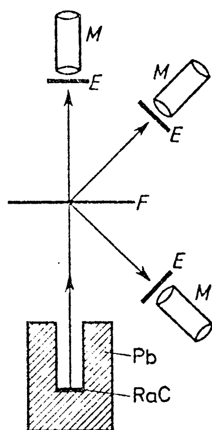
Rozważania rozpoczniemy od pierwszych lat bieżącego stulecia. Jest to już okres po odkryciu elektronu i po odkryciu zjawiska promieniotwórczości naturalnej. Modelowe przedstawienie atomu w tym okresie odpowiada kuli. W całej objętości tej kuli jest równomiernie rozłożony ładunek dodatni i masa. Wewnątrz kuli rozmieszczone są elektrony z zachowaniem pewnej symetrii. Elektrony mogą wykonywać ruchy drgające dokoła swoich położeń równowagi. Liczba elektronów w atomie jest taka, że łączny ładunek elektryczny atomu wynosi zero. Atom, jako całość, jest więc elektrycznie obojętny. Taki model wystarczał do wytłumaczenia znanego w tym czasie tzw. *normalnego zjawiska Zeemana*, polegającego na rozszczepieniu linii widmowych w polu magnetycznym. Nie tłumaczył jednak tworzenia się serii widmowych i zawodził przy wyjaśnianiu wyników badań przechodzenia przez materię promieni katodowych i cząstek α.

Już badania Lenarda nad przechodzeniem promieni katodowych przez cienkie płytki metalowe doprowadziły go do wniosku, że masa nie jest w atomie rozłożona równomiernie. Potwierdzenie tego wniosku przyniosły późniejsze badania Rutherforda, Geigera, Marsdena, Wilsona i innych nad przejściem cząstek α przez materię. Nie wnikając w to, jak w owym czasie wyobrażano sobie budowę cząstki α, przypominamy, że obecnie traktujemy cząstkę α wysyłaną przez pierwiastki promieniotwórcze jako cząstkę złożona z dwóch protonów i dwóch neutronów. Proton ma masę $1{,}6725 \cdot 10^{-27}$ kg i ładunek elementarny dodatni $1{,}602 \cdot 10^{-19}$ C. Neutron jest cząstką elektrycznie obojętną o masie $1{,}6748 \cdot 10^{-27}$ kg bardzo zbliżonej do masy protonu. Cząstka α odpowiada pod względem budowy dwuwartościowemu jonowi helu. Masa cząstki α wynosi $6{,}646 \cdot 10^{-27}$ kg i jest około 7300 razy większa od masy elektronu.

Rysunek 30.2 przedstawia schemat doświadczenia Rutherforda, Geigera i Marsdena dotyczącego rozpraszania cząstek α. Wąska wiązka promieni α wychodzi z preparatu promieniotwórczego, np. z RaC, umieszczonego na dnie wydrążenia w ołowianym bloku Pb, i zostaje skierowana na bardzo cienką (rzędu setnych części milimetra) złotą folię *F*. Dalszy bieg cząstek α badany jest metodą scyntylacji, czyli iskrzeń, jakie cząstki α wywołują na fluoryzującym ekranie *E*. Ekran ten obserwowany jest przez mikros-

kop M. Licząc występujące iskrzenia określamy liczbę cząstek α dobiegających do ekranu. Zmieniając kąt ustawienia ekranu można ustalić rozkład przestrzenny cząstek α rozpraszanych przez folię. Różnym kątom odchylenia od pierwotnego biegu odpowiadają różne liczby dobiegających do ekranu cząstek α.

Badania tego rodzaju wykazały, że większość cząstek biegnie z małym odchyleniem od 2° do 4°, ale przeciętnie na 1000 cząstek odchylonych od pierwotnego biegu o 30° przypada 9 odchylonych o 90°, a 3 odchylone pod kątem 150°. Doświadczenia powtarzano wielokrotnie z różnymi foliami, uzyskując podobne wyniki.



Rys. 30.2

Zestawienie tych wyników doprowadziło Rutherforda do następującego modelu budowy atomu. W każdym atomie można wyodrębnić część środkową, zwaną *jądrem*, w której skupia się całkowity ładunek dodatni i prawie cała masa atomu. Rozmiary tej części są bardzo małe w porównaniu z rozmiarami atomu jako całości, np. promień atomu złota jest rzędu 10^{-10} m, a promień jądra szacuje się na 10^{-14} m. Dokoła jądra po torach kołowych krążą elektrony (promień elektronu jest rzędu 10^{-15} m). Liczba elektronów jest taka, że ich łączny ładunek ujemny neutralizuje ładunek dodatni jądra. Jak widać z tego modelowego ujęcia, wewnątrz atomu jest stosunkowo bardzo dużo wolnej przestrzeni.

Stosując taki model atomu łatwo wyjaśnić zachowanie się cząstek α w opisanych wyżej doświadczeniach. Cząstka α może przebiegać w różnych odległościach od jądra. Przy stosunkowo dużej odległości między jądrem i cząstką α siła kulombowskiego odpychania ładunków jednoimiennych będzie niewielka. Taka cząstka wyminie jądro, nieznacznie tylko zmieniając kierunek swego ruchu.

Jeśli przedłużenie pierwotnego toru cząstki α wypadnie blisko jądra, odchylenie będzie odpowiednio większe. Gdyby przedłużenie toru trafiało wprost w jądro atomu, to nawet wtedy nie doszłoby do mechanicznego zderzenia cząstki α z jądrem. Przy dostatecznym zbliżeniu cząstki α do jądra siła wzajemnego odpychania byłaby tak duża, że cząstka α zostałaby zahamowana w swym biegu i odrzucona z powrotem w tym kierunku, z jakiego nadbiegła.

Stosunkowo rzadko zdarzające się odrzuty cząstek α wstecz wynikają z małych rozmiarów jądra atomu. Prawdopodobieństwo trafienia małej co do rozmiarów cząstki α w małe jądro atomowe jest bardzo niewielkie.

Oddziaływanie elektronów na bieg cząstek α jest znikome, zarówno z powodu ich małej masy, jak i małego ładunku (w porównaniu z ładunkiem jąder ciężkich pierwiastków). Raczej cząstka α, przelatując obok przeszło 7000 razy lżejszego elektronu, będzie miała wpływ na jego ruch.

Rysunek 30.3 przedstawia kilka torów cząstki α przelatującej w różnych odległościach od jądra J. Cząstka na torze I, dalekim od jądra, biegnie prawie prostoliniowo. Tory II i III, bliższe jądra, mają przebieg hiperboliczny. Interesujący jest bieg cząstki IV, kierującej się wprost na jądro. W miarę zbliżania się do jądra siła odpychania elektrycznego rośnie i, jak już mówiliśmy, do zderzenia mechanicznego nie dochodzi.

Nasuwa się pytanie, na jaką odległość cząstka α zbliży się do jądra. Zależeć to będzie od początkowej energii kinetycznej cząstki α. Kosztem energii kinetycznej powstaje energia potencjalna (elektryczna) rosnąca w miarę zbliżania się cząstki α do jądra. Jeśli jądro należy do atomu o liczbie porządkowej Z, to jego ładunek dodatni wynosi Ze, gdzie e oznacza ładunek elementarny. Potencjał w odległości r od takiego ładunku wynosi $\frac{1}{4\pi\varepsilon_0}\frac{Ze}{r}$, a energia potencjalna cząstki α o ładunku $2e$ w tejże odległości wynosi

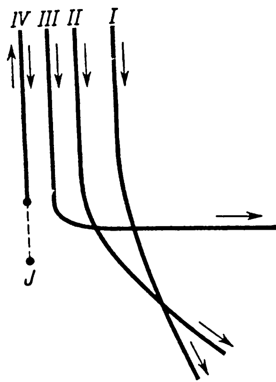
$$\frac{1}{4\pi\varepsilon_0}\frac{Ze}{r}2e,$$

gdzie ε_0 oznacza przenikalność elektryczną próżni.

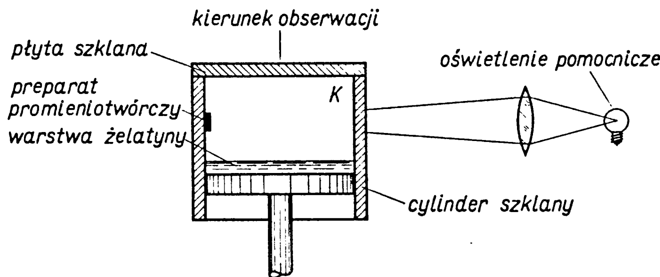
Cząstka α o masie m zbliży się do jądra na taką odległość r, która wynika z równości

$$\frac{mv^2}{2}=\frac{1}{4\pi\varepsilon_0}\frac{2Ze^2}{r}.$$

Dla folii z takich metali, jak złoto, srebro i platyna odległość maksymalnego zbliżenia cząstki α do jądra jest rzędu 10^{-14} m.



Rys. 30.3



Rys. 30.4

Rutherford wyprowadził też wzory wiążące ładunek jądra Ze z najbardziej prawdopodobnym kątem odchylenia cząstki α od pierwotnego biegu.

Wilson w roku 1912 opracował metodę bezpośredniego badania torów cząstek α. Zapoznamy się z zasadą działania tzw. *komory Wilsona*, za pomocą której można fotografować tory różnych cząstek wywołujących jonizację, między innymi tory cząstek α.

Przy badaniu torów cząstek w komorze Wilsona wykorzystano zjawisko skraplania się pary na jonach wytwarzanych wzdłuż torów przelatujących cząstek. Rysunek 30.4 przedstawia schemat urządzenia. Komora K, zawierająca sprężone powietrze, zamknięta jest od dołu tłokiem T, na którym znajduje się warstewka wilgotnej żelatyny. Dzięki temu mamy pewność, że w komorze znajduje się para nasycona wody w danej temperaturze. W czasie gwałtownego przesunięcia tłoka w dół następuje rozprężenie, któremu towarzyszy obniżenie temperatury. W nowej, niższej temperaturze para wodna staje

się przesycona. Wtedy nadmiar pary powinien ulec skropleniu. Żeby jednak przesyceniu pary mogło towarzyszyć jej skroplenie, muszą istnieć w komorze Wilsona zarodki kondensacji pary w postaci cząstek kurzu lub jonów. Jeśli nie dopuścimy do tego, aby w komorze był kurz, to jedynymi zarodkami kondensacji będą jony.

Wprowadźmy do komory Wilsona źródło cząstek α w postaci odpowiedniego preparatu promieniotwórczego. Wzdłuż toru cząstki wybiegającej z preparatu tworzyć się będą jony. Z rozmaitych preparatów cząstki α wylatują z różnymi prędkościami (rzędu 10^7 m/s), a więc z różnymi energiami kinetycznymi. Energia kinetyczna cząstki α przetwarza się na energię jonizacji napotykanych cząsteczek. Cząstki α można charakteryzować podając ich zasięg w powietrzu w warunkach normalnych. *Zasięg cząstki α* jest to odległość, wzdłuż której cząstka α jonizuje napotkane atomy lub cząsteczki. Na przykład cząstka α o zasięgu około 7 cm w powietrzu w warunkach normalnych jonizuje około 100 000 napotkanych atomów, tzn. wybija z nich elektrony. Atom pozbawiony elektronu stanowi jon dodatni, atom obojętny, do którego dołączy się elektron, stanowi jon ujemny. Jeśli taka cząstka α przebiegnie przez komorę Wilsona niezwłocznie po dokonaniu rozprężenia, to 100 000 par jonów utworzonych na jej drodze będzie stanowiło gęsty zbiór ośrodków kondensacji. Cały tor cząstki α zaznaczy się jako ciąg drobniutkich kropelek mgły. Robiąc w tym samym momencie zdjęcie fotograficzne otrzymamy wyraźny tor cząstki α.

Olbrzymia większość torów cząstek α w powietrzu ma przebieg prostoliniowy. Niektóre tory wykazują odchylenia przy końcu, wtedy gdy już energia biegnącej cząstki jest stosunkowo niewielka.

30.3. Model atomu Rutherforda–Bohra

Zestawmy raz jeszcze wnioski Rutherforda dotyczące budowy atomu. Atom składa się z części wewnętrznej, zwanej jądrem, skupiającej prawie całą masę, cały ładunek dodatni, i mającej promień rzędu 10^{-14} m. Dokoła jądra krążą elektrony w takiej liczbie, że w stanie normalnym atom jako całość jest elektrycznie obojętny.

Taki obraz budowy atomu nie był jednak zadowalający. Nie tłumaczył ani trwałości budowy atomu, ani nie wyjaśniał tworzenia się charakterystycznych widm liniowych.

Braki te widział między innymi fizyk duński Niels Bohr, który przez pewien czas pracował w laboratorium Rutherforda. Z jego rozważań nad budową atomu powstaje nowa koncepcja modelowa, tzw. model *atomu Rutherforda–Bohra*.

Z punktu widzenia elektrodynamiki klasycznej jądro wraz z krążącymi dokoła niego elektronami musiałoby stanowić twór nietrwały. Z ruchem elektronów dokoła jądra musiałoby się wiązać wypromieniowywanie energii w postaci fal elektromagnetycznych. Równocześnie malałaby energia kinetyczna elektronów, a co za tym idzie — ich prędkość i promień orbity. Okres ruchu elektronów na orbicie malałby, a więc wzrastałaby częstotliwość wysyłanej fali. Po pewnym czasie krążący elektron w swoim ruchu po spirali spadałby na jądro, a więc atom nie miałby budowy trwałej. W czasie krążenia elektronu wysyłane byłyby fale elektromagnetyczne o częstotliwościach zmieniających

się w sposób ciągły, czyli otrzymalibyśmy widmo ciągłe, a nie widmo liniowe, charakterystyczne dla atomów danego pierwiastka.

Bohr decyduje się na bardzo radykalne posunięcie, a mianowicie *odrzuca stosowalność zasad elektrodynamiki klasycznej do zjawisk wewnątrzatomowych.* Na miejsce odrzuconych zasad elektrodynamiki klasycznej Bohr wprowadza nowe założenia, które nazywamy *postulatami Bohra.* Zgodnie z *pierwszym postulatem Bohra,* w atomie istnieją pewne dozwolone, *stacjonarne tory* (*orbity*), po których elektron może krążyć bez wypromieniowywania energii. Orbity te Bohr traktuje jako tory kołowe. Każdemu z dozwolonych torów w atomie narzuca Bohr następujący warunek kwantowy:

$$r_n m v_n = n\frac{h}{2\pi}, \tag{30.4}$$

gdzie $n = 1, 2, 3, \ldots$

Zgodnie z tym warunkiem *moment pędu $r_n m v_n$ elektronu na n-tej orbicie jest całkowitą wielokrotnością stałej Plancka h dzielonej przez 2π* ($h/2\pi$ oznacza się często $\hbar$ i nazywa się *h kreślonym*).

Drugi postulat Bohra przypisuje elektronowi na n-tej orbicie stacjonarnej ściśle określoną wartość energii E_n. Torom bliższym jądra odpowiada mniejsza energia. Tłumaczy się to tym, że do odsunięcia elektronów na większą odległość od jądra potrzebna jest większa praca pokonania przyciągania elektrycznego.

W stanie normalnym atomu elektrony znajdują się na najbliższych jądra orbitach stacjonarnych. *Wzbudzenie atomu* wiąże się z przeniesieniem elektronu na orbitę wyższą, położoną dalej od jądra, a więc odpowiadającą większej energii. Wzbudzenie atomu jest zatem skutkiem pochłonięcia porcji energii równej różnicy wartości energii toru końcowego E_k i toru początkowego E_p. Ta wartość energii $E_k - E_p$ może być dostarczona w rozmaity sposób, np. przez zderzenie z atomami, przez zderzenie z pędzącym elektronem, przez pochłonięcie odpowiedniego kwantu promieniowania itp. Wzbudzenie może być tak silne, że elektron zostanie odrzucony poza sferę działania jądra. Wtedy mówimy o *jonizacji atomu.*

Atom zdolny jest do wysyłania promieniowania tylko wtedy, gdy został przedtem wzbudzony. *Emisja promieniowania* wiąże się z przeskokiem elektronu z orbity wyższej na niższą. Towarzysząca temu przeskokowi zmiana energii atomu $E_k - E_p$ zostaje zużyta całkowicie na wytworzenie energii kwantu promieniowania. Jest to treść *trzeciego postulatu Bohra*:

$$E_k - E_p = h\nu.$$

Znaczenie symboli jest nam już znane. Jak widzimy, częstotliwość wysyłanego promieniowania ν zależy jedynie od różnicy wartości energii $E_k - E_p$. Nie jest to częstotliwość krążenia elektronu po tej lub innej orbicie elektronowej. A więc *według Bohra zarówno pochłanianie, jak wysyłanie energii przez atom może odbywać się tylko ściśle określonymi porcjami, kwantami.*

Przeanalizujemy obecnie dokładniej, do jakich wniosków doprowadza zastosowanie postulatów Bohra do atomu wodoru.

Atom wodoru składa się z jądra, czyli protonu o pojedynczym ładunku elementarnym e (dodatnim), i z jednego elektronu, mającego ładunek elementarny $-e$ (ujemny). Krążenie elektronu dokoła jądra po torze kołowym odbywa się pod działaniem siły dośrodkowej, wywołanej przyciąganiem elektrycznym jądra:

$$\frac{mv_n^2}{r_n} = \frac{1}{4\pi\varepsilon_0}\frac{e^2}{r_n^2}. \tag{30.5}$$

W równaniu tym mamy dwie niewiadome: prędkość v_n elektronu na danej orbicie i jej promień r_n.

Rozwiązujemy układ równań (30.4) i (30.5) względem r_n. Z równania (30.4) znajdujemy

$$v_n = \frac{nh}{2\pi r_n m},$$

a po podstawieniu do równania (30.5) otrzymujemy

$$r_n = \frac{\varepsilon_0 n^2 h^2}{\pi m e^2}. \tag{30.6}$$

Podstawiając zamiast n kolejno wartości 1, 2, 3, ..., zamiast m — masę elektronu, e — ładunek elementarny, h — stałą Plancka, otrzymamy promienie kolejnych dozwolonych orbit w atomie wodoru. Dla orbity najbliższej jądra otrzymujemy

$$r_1 \approx 0{,}5\ \text{Å} = 0{,}5\cdot 10^{-10}\ \text{m}.$$

Promienie dalszych orbit rosną, jak widać ze wzoru, proporcjonalnie do kwadratu liczby n.

Z tego samego układu równań otrzymujemy wyrażenie na v_n:

$$v_n = \frac{nh\pi m e^2}{2\pi m n^2 h^2 \varepsilon_0},$$

czyli

$$v_n = \frac{e^2}{2nh\varepsilon_0}. \tag{30.7}$$

Po podstawieniu wartości liczbowych znajdujemy, że prędkość elektronu na orbicie pierwszej, najbliższej jądra, równa się $\frac{1}{137}$ prędkości światła. Na dalszych orbitach prędkości maleją proporcjonalnie do wielkości n.

W atomie wodoru w stanie normalnym (niewzbudzonym) elektron znajduje się na orbicie najbliższej jądra. Podczas wzbudzenia elektrony w poszczególnych atomach przechodzą na orbity wyższe — drugą, trzecią itd., zależnie od pobranej porcji energii. W olbrzymim skupisku atomów, z jakim mamy zazwyczaj do czynienia, występują po wzbudzeniu tysiące atomów wodoru mających elektrony na orbicie drugiej, trzeciej i wyższych. Stan wzbudzenia jest krótkotrwały. Przy przeskokach z orbit wyższych na niższe wysyłane są określone porcje energii $h\nu_1$, $h\nu_2$ itd. Wobec jednoczesnego dokonywania się jednakowych przeskoków w tysiącach atomów każdy kwant energii $h\nu$ jest reprezentowany wielokrotnie w emitowanym promieniowaniu. Przy dostatecznie dużej liczbie emitowanych kwantów $h\nu$ otrzymujemy takie natężenie światła, na które

już reaguje nasze oko lub klisza fotograficzna. Gdyby elektrony we wszystkich atomach wodoru zostały wzbudzone w ten sposób, że wykonałyby przeskok z orbity pierwszej na drugą, to przy emisji otrzymalibyśmy światło monochromatyczne o jednej, określonej częstotliwości. Po przejściu promieniowania przez odpowiedni spektrograf otrzymalibyśmy tylko jedną linię widmową: byłaby to linia promieniowania niewidzialnego, leżąca w nadfiolecie.

Gdyby podczas wzbudzenia elektrony we wszystkich atomach zostały przeniesione na orbitę trzecią, licząc od jądra, to przeskoki odpowiadające emisji mogłyby być trzech rodzajów. W niektórych atomach elektron mógłby wracać bezpośrednio z poziomu trzeciego do pierwszego, emitując promieniowanie o częstotliwości ν''', spełniającej równanie

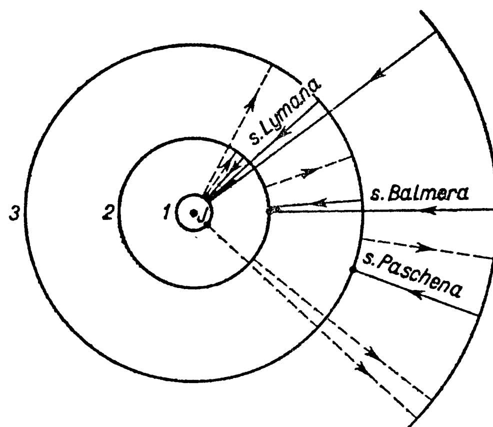
$$E_3 - E_1 = h\nu'''.$$

W innych atomach przeskok powrotny mógłby się dokonywać „na raty", najpierw z poziomu trzeciego na drugi przy emisji spełniającej równanie

$$E_3 - E_2 = h\nu''$$

i następnie z drugiego na pierwszy

$$E_2 - E_1 = h\nu'.$$



Rys. 30.5

Tego rodzaju przeskoki, odbywające się jednocześnie w bardzo wielu atomach, dałyby w badaniach spektrograficznych linie widmowe o określonych częstotliwościach ν''', ν'', i ν' i o określonych długościach fal w powietrzu. Rozważanie innych możliwości wzbudzenia i powrotu wzbudzonego atomu wodoru do stanu normalnego doprowadziłoby do dalszych linii widmowych wodoru. W § 30.1 mówiliśmy już, że linie widmowe wodoru zostały zbadane doświadczalnie, ich częstotliwości i długości fal wymierzone, a one same zostały stopniowo zgrupowane przez badaczy w kilka serii.

Stosując model Bohra można wyjaśnić mechanizm tworzenia się serii wodorowych (rys. 30.5). Na serię Lymana składają się takie linie, którym odpowiadają przeskoki

elektronowe z dowolnych orbit wyższych na orbitę pierwszą. Serię Balmera tworzą linie powstające dzięki przeskokom elektronowym z dowolnych poziomów wyższych na drugi. Podobnie, przeskoki kończące się na orbicie trzeciej prowadzą do linii serii Paschena, na orbicie czwartej — do linii serii Bracketta, itd.

Na rysunku 30.5 przedstawione są: 1) przeskoki elektronowe prowadzące do wzbudzenia atomu wodoru, a więc związane z absorpcją energii (linie kreskowane na rysunku); 2) przeskoki elektronowe odpowiadające emisji promieniowania związanej z powstawaniem linii odpowiednich serii (linie ciągłe na rysunku).

Pamiętamy, że linie widmowe wodoru należące do dowolnej serii można podporządkować wzorowi (30.2):

$$\frac{1}{\lambda} = R_H\left(\frac{1}{n^2} - \frac{1}{s^2}\right),$$

gdzie $n = 1, 2, 3, \ldots$, a $s = n+1, n+2, n+3, \ldots$

Nasuwa się pytanie, czy stosując postulaty Bohra i opierając się na narzuconych przez niego warunkach kwantowych, można wyprowadzić teoretycznie taki sam wzór zgodny z doświadczeniem.

Rozważania teoretyczne Bohra uwieńczone zostały powodzeniem. Udało mu się na podstawie założeń kwantowych wyprowadzić wzory określające energię elektronów na poszczególnych orbitach. Od wartości energii na poszczególnych orbitach łatwo można przejść do częstotliwości promieniowania odpowiadających przeskokom elektronowym, a stąd już jeden krok do długości fali lub też do jej odwrotności $\tilde{\nu}$, czyli liczby falowej. Prześledzimy dokładniej rozumowanie Bohra. Całkowita energia elektronu na orbicie n składa się z energii kinetycznej i energii potencjalnej. Energię kinetyczną można ogólnie wyrazić następująco:

$$E_k = \frac{mv^2}{2} = \frac{1}{2}\frac{mv^2}{r}r.$$

Czynnik mv^2/r przedstawia siłę dośrodkową, a więc może być zastąpiony wyrażeniem $e^2/4\pi\varepsilon_0 r^2$. A zatem energia kinetyczna na orbicie n

$$E_k = \frac{1}{2}\frac{e^2}{4\pi\varepsilon_0 r_n}.$$

Energia potencjalna może być wyznaczona w stosunku do określonego poziomu. Załóżmy, że wartość energii potencjalnej w odległości nieskończonej od jądra wynosi 0. W punkcie odległym od jądra o r energia będzie mniejsza o wartość pracy potrzebnej do przeniesienia ładunku $-e$ z nieskończoności do odległości r, będzie zatem ujemna. Z wzorów elektrostatycznych wiadomo, że praca ta równa się iloczynowi przenoszonego ładunku i różnicy potencjałów w punkcie początkowym i końcowym drogi, czyli

$$E_p = -[-e(V_\infty - V_r)],$$

$$E_p = -e(V_r - V_\infty).$$

Ale potencjał w nieskończoności równa się zeru, a potencjał w odległości r wynosi $e/4\pi\varepsilon_0 r$. A zatem

$$E_p = -\frac{e^2}{4\pi\varepsilon_0 r}.$$

Całkowita energia elektronu na orbicie n wynosi

$$E_n = \left(\frac{1}{2}\frac{e^2}{r_n} - \frac{e^2}{r_n}\right)\frac{1}{4\pi\varepsilon_0},$$

czyli

$$E_n = -\frac{1}{2}\frac{e^2}{r_n}\frac{1}{4\pi\varepsilon_0}. \tag{30.8}$$

Uwzględniając wartość r_n (wzór (30.6)) otrzymujemy

$$E_n = -\frac{e^2\pi me^2}{8\pi\varepsilon_0^2 h^2 n^2},$$
$$E_n = -\frac{me^4}{8\varepsilon_0^2 n^2 h^2}. \tag{30.9}$$

Analogicznie energia elektronu na orbicie s wynosi

$$E_s = -\frac{me^4}{8\varepsilon_0^2 s^2 h^2}.$$

Energia emitowanego kwantu równa się różnicy obu energii:

$$h\nu = E_s - E_n = \frac{me^4}{8\varepsilon_0^2 h^2}\left(\frac{1}{n^2} - \frac{1}{s^2}\right). \tag{30.10}$$

Uwzględniając zależność

$$\nu = \frac{c}{\lambda}$$

znajdujemy

$$\frac{1}{\lambda} = \frac{me^4}{8\varepsilon_0^2 h^3 c}\left(\frac{1}{n^2} - \frac{1}{s^2}\right). \tag{30.11}$$

Uderzająca jest analogia tego wzoru do wzoru Balmera (30.2). Stała Rydberga dla wodoru powinna być równa

$$R_{\mathrm{H}} = \frac{me^4}{8\varepsilon_0^2 h^3 c}. \tag{30.12}$$

(W § 30.5 wykażemy, że słuszniejsze byłoby wprowadzenie symbolu R_∞ zamiast R_{H}.) Okazuje się, że różnica wartości R_{H}, obliczonej na podstawie tego wzoru i wyznaczonej doświadczalnie, jest mniejsza niż 1%.

Wzór otrzymany przez Bohra jest ogólniejszy niż wzór Balmera, obejmuje bowiem wszystkie znane serie wodorowe. Należy podkreślić, że przed wysunięciem koncepcji modelu Bohra (tzn. przed rokiem 1913) znane były tylko dwie serie wodorowe. Tym razem teoria wyprzedziła doświadczenie.

30.4. Poziomy energetyczne

Istnieje wygodny sposób przedstawiania mechanizmu promieniowania atomu za pomocą schematycznego układu poziomów energetycznych w atomie. Rozpatrzymy taki schemat na przykładzie atomu wodoru. Orbitom stacjonarnym w atomie wodoru przypisaliśmy określone wartości energii zależne, zgodnie z wzorem (30.9), od liczby n, zwanej *liczbą kwantową główną*. Stan podstawowy ($n = 1$) ma energię

$$E_1 = -\frac{me^4}{8\varepsilon_0^2 h^2}\frac{1}{1^2},$$

stan drugi ($n = 2$) ma energię

$$E_2 = -\frac{me^4}{8\varepsilon_0^2 h^2}\frac{1}{2^2},$$

stan trzeci ($n = 3$) ma energię

$$E_3 = -\frac{me^4}{8\varepsilon_0^2 h^2}\frac{1}{3^2}$$

itd. Możemy tak dobrać układ linii poziomych, aby ich odległości od linii najwyższej N, odpowiadającej energii zerowej, wyrażały w pewnej skali podane wyżej wartości E_1, E_2, E_3, ... itd. (rys. 30.6).

Przy takim założeniu można uważać, że linia najniższa odpowiada stanowi energetycznemu $n = 1$, linia druga od dołu — stanowi energetycznemu $n = 2$ itd. Linie tego układu reprezentują poziomy energetyczne atomu. Na rysunku widać wyraźnie zagęszczanie się poziomów energetycznych w miarę przechodzenia do stanów energetycznych wyższych. Gdyby skala była dokładnie zachowana, to zagęszczanie się linii byłoby jeszcze wyraźniejsze. Przy kilku niższych poziomach energetycznych z prawej strony podana jest wartość energii w dżulach. Wartości $n = \infty$ odpowiada energia zerowa.

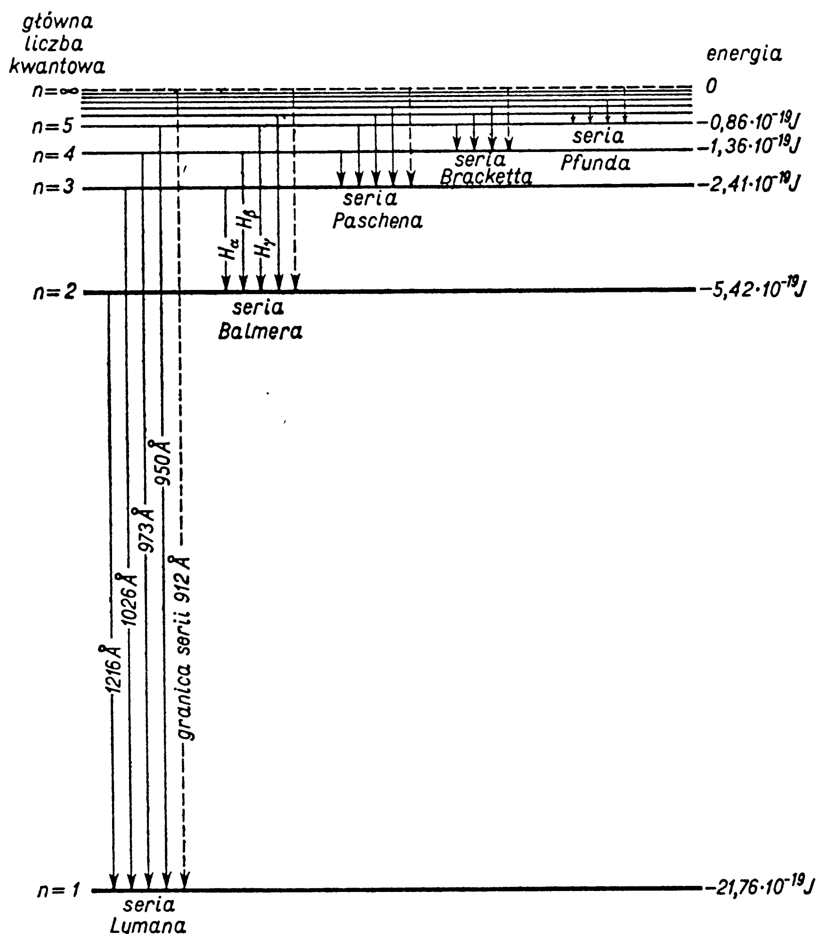
Przejściu elektronu z jednego poziomu energetycznego na drugi odpowiada różnica energii ΔE, pochłonięta przez atom w przypadku jego wzbudzenia lub wypromieniowana w postaci kwantu $h\nu$. Na przykład różnica

$$\left(-\frac{me^4}{8\varepsilon_0^2 h^2}\frac{1}{2^2}\right)-\left(-\frac{me^4}{8\varepsilon_0^2 h^2}\frac{1}{1^2}\right) = h\nu$$

odpowiada przejściu z poziomu $n = 2$ na $n = 1$ i wysłaniu kwantu odpowiadającego pierwszej linii Lymana. Analogiczne przejście z $n = 3$ do $n = 1$ odpowiada drugiej linii serii Lymana itd.

Przejścia z poziomów niższych na wyższe, a więc odwrotne do zaznaczonych na rys. 30.6, przedstawiałyby akty absorpcji energii.

Aby pogłębić nieco wiadomości dotyczące emisji i absorpcji promieniowania przez wodór, zastanówmy się jeszcze nad możliwościami pochłaniania i emisji kwantów promieniowania odpowiadających np. liniom serii Balmera.



Rys. 30.6

Wiemy już, że przypisując elektronom możliwość odbywania ruchu po płaskich orbitach kołowych dokoła centralnego nieruchomego jądra Bohr wyprowadził wzory na promień orbity (30.6), na prędkość elektronu na orbicie (30.7) oraz — co w tej chwili jest najważniejsze — wzór na energię elektronu na orbicie (30.9). W tym ostatnim wzorze widoczna jest zależność od liczby n, która przyjmuje wyłącznie wartości całkowite od 1 do ∞. A zatem i energia nie zmienia się w sposób ciągły, lecz przyjmuje tylko pewne określone — *dyskretne* — wartości, czyli innymi słowy, *energia jest skwantowana.*

Jeżeli weźmiemy pod uwagę atom wodoru w stanie normalnym, to jedyny elektron musi zajmować miejsce na orbicie o $n = 1$, której odpowiada minimalna energia. Do wzbudzenia tego atomu potrzebna jest porcja energii z zewnątrz. Nie może to być jednak dowolna porcja energii. Jak widać z rys. 30.6, minimalna energia potrzebna do wzbudzenia atomu wodoru w stanie normalnym wynosi $16{,}34 \cdot 10^{-19}$ J (czyli 10,21 eV). Jest to energia większa od energii kwantu odpowiadającego dowolnej linii serii Balmera. A zatem wodór niewzbudzony jest całkowicie przezroczysty dla kwantów promieniowania odpowiadających liniom Balmera.

Inaczej się sprawa przedstawia z wodorem wzbudzonym. Wśród atomów wzbudzonych wodoru (np. dzięki zderzeniom z innymi atomami czy też dzięki zderzeniom z elektronami w czasie wyładowań elektrycznych w rozrzedzonym wodorze w rurce Geisslera) na pewno są takie, w których elektrony znajdują się na orbicie drugiej. Takie atomy są przystosowane do absorpcji kwantów odpowiadających liniom Balmera. Absorpcji tych kwantów towarzyszą przeskoki elektronów na orbity 3, 4, 5 lub wyższe. Innymi słowy, wodór wzbudzony jest nieprzezroczysty dla kwantów promieniowania serii Balmera: nie przepuszcza takich kwantów, lecz je pochłania. Tym właśnie tłumaczy się istnienie w widmie absorpcyjnym Słońca linii Fraunhofera, odpowiadających co do długości liniom serii Balmera. W skład atmosfery Słońca, dzięki jej wysokiej temperaturze, wchodzą atomy wodoru nie w stanie normalnym, lecz w stanie wzbudzonym. Zgodnie z tym, cośmy wyżej powiedzieli, atomy te absorbują z promieniowania ciągłego, idącego od powierzchni Słońca, te kwanty, które odpowiadają liniom serii Balmera.

Rysunek 30.6 pozwala też na wyjaśnienie pojęcia energii jonizacji. Dostarczenie niewzbudzonemu elektronowi w atomie wodoru energii równej $21{,}76 \cdot 10^{-19}$ J (czyli 13,6 eV) powoduje jego przejście do stanu $n = \infty$. W tych warunkach elektron staje się całkowicie niezależny od macierzystego atomu, który przekształca się w jon. Wymieniona energia jest właśnie *energią jonizacji wodoru.*

Jak już podkreślaliśmy, wzbudzenie atomu może nastąpić m.in. wskutek zderzenia jego elektronu z rozpędzonym elektronem swobodnym. Energia kinetyczna bombardującego elektronu zamienia się w energię wzbudzenia. Muszą być jednak przy tym zachowane warunki kwantowe i w związku z tym należy rozróżnić następujące przypadki. Jeśli energia kinetyczna bombardującego elektronu jest mniejsza od energii przejścia elektronu bombardowanego na wyższy poziom energetyczny, to w ogóle nie pochłania on energii i *zderzenie* jest *doskonale elastyczne* (sprężyste). Jeśli energia bombardującego elektronu jest większa lub równa energii przejścia do najbliższego poziomu, lecz mniejsza od energii jonizacji, to elektron orbitalny może pobrać część lub całość tej energii, przechodząc do stanu energetycznie wyższego (w dalszej konsekwencji wystąpi emisja promieniowania). Ten rodzaj zderzenia jest *zderzeniem nieelastycznym* (*niesprężystym*).

Gdy energia bombardującego elektronu jest większa lub co najmniej równa energii jonizacji, zderzenia elektronów prowadzą do wybicia elektronu orbitalnego z atomu (jonizacji).

30.5. Jony wodoropodobne

Powróćmy jeszcze do zagadnień związanych ze spektroskopią. Wiadomo, że przez utracenie jednego lub więcej elektronów atom przekształca się w odpowiedni jon. Jonizacja może wystąpić np. w wysokiej temperaturze iskry elektrycznej. Dlatego też widma zjonizowanych pierwiastków nazywamy często *widmami iskrowymi*, w odróżnieniu od *widm łukowych* atomów obojętnych. W spektroskopii obowiązuje tzw. *spektroskopowe prawo przesunięć*. Zgodnie z tym prawem serie linii iskrowych n-krotnie zjonizowanych atomów pierwiastka o liczbie atomowej Z są analogiczne do serii linii łukowych obojętnych atomów pierwiastka o liczbie atomowej $Z-n$. Stąd wniosek, że w miarę przechodzenia do pierwiastków coraz cięższych układ elektronów w warstwach położonych bliżej jądra powtarza się bez zmian.

Rozpatrzymy prawo przesunięć w odnięsieniu do konkretnych przypadków. Weźmy pod uwagę pojedynczo zjonizowany atom helu ($Z = 2$) lub podwójnie zjonizowany atom litu ($Z = 3$). Zgodnie z prawem przesunięć widma tych zjonizowanych pierwiastków powinny odpowiadać co do charakteru widmu wodoru. Doświadczenie potwierdza to zarówno w przypadku jonów wyżej wymienionych, jak i w przypadku berylu potrójnie zjonizowanego, boru zjonizowanego poczwórnie i węgla zjonizowanego pięciokrotnie. Mówimy w tych przypadkach o *widmach wodoropodobnych*. Należy jednak podkreślić, że odpowiednie długości fal w poszczególnych seriach widm wodoropodobnych są mniejsze niż w seriach wodorowych, gdyż we wzorach na serie widmowe pierwiastków cięższych występuje dodatkowy czynnik Z^2, a mianowicie

$$\frac{1}{\lambda} = Z^2 R\left(\frac{1}{n^2} - \frac{1}{s^2}\right). \tag{30.13}$$

P y t a n i e. 1. Prześledzić wyprowadzenie wzoru Bohra dla jonu wodoropodobnego o liczbie atomowej Z (tzn. udowodnić wzór (30.13)). 2. Na podstawie wzoru (30.13) wykazać, że co drugi prążek (parzysty) każdej parzystej serii pojedynczo zjonizowanego helu miałby tę samą długość fali co kolejne prążki poprzedzającej serii wodorowej, gdyby stałe R dla wodoru i helu były jednakowe (wziąć dla przykładu λ odpowiadające: a) $n = 2$ i $s = 4$ dla zjonizowanego helu oraz $n = 1$ i $s = 2$ dla wodoru, b) $n = 2$, $s = 6$ dla zjonizowanego helu oraz $n = 1$, $s = 3$ dla wodoru).

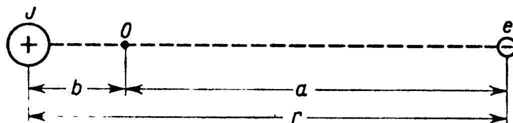
Uwzględnione w pytaniu drugim zastrzeżenie co do ewentualnych różnic stałych R dla wodoru i helu (ogólniej mówiąc — zależność R od masy jądra atomu) wymaga dokładniejszego omówienia. Całość przytoczonych rozważań matematycznych prowadzących do wzoru (30.12) oparta była na założeniu, że jądro jest nieruchome, a tylko elektron krąży dokoła niego, czyli, innymi słowy, że masa jądra jest nieskończenie wielka w porównaniu z masą elektronu. Otrzymaną przy tym założeniu wartość R_H (wzór (30.12)) słuszniej byłoby oznaczać R_∞. Odrzucenie wymienionego założenia wymaga uwzględnienia ruchu obrotowego układu jądro-elektron dokoła wspólnego środka masy (rys. 30.7). Odległości a i b elektronu i jądra od wspólnego środka masy O wynikają bezpośrednio z równań mechaniki klasycznej:

$$a = r\frac{M}{m+M}, \quad b = r\frac{m}{m+M}$$

lub
$$a = r\frac{1}{1+\frac{m}{M}}, \quad b = r\frac{\frac{m}{M}}{1+\frac{m}{M}}.$$

Uwzględnienie ruchu jądra wymaga wprowadzenia poprawki zależnej od stosunku masy m elektronu do masy M jądra. (Dla wodoru stosunek ten wynosi $\frac{1}{1836,5}$, dla helu jest jeszcze cztery razy mniejszy.) Uwzględnienie tych poprawek prowadzi do wzorów:
$$R_H = \frac{R_\infty}{1+\frac{m}{M_H}}, \quad R_{He} = \frac{R_\infty}{1+\frac{m}{M_{He}}}.$$

Liczbowe wartości R_H i R_{He} wynoszą:
$$R_H = 109\,677\,58\ \mathrm{m}^{-1}, \quad R_{He} = 109\,722\,27\ \mathrm{m}^{-1}.$$



Rys. 30.7

30.6. Warunki kwantowe Sommerfelda

Gdybyśmy poprzestali na podaniu powyższych wiadomości, stworzylibyśmy fałszywy obraz rzeczywistości. Pozostalibyśmy pod błędnym wrażeniem, że model Bohra tłumaczy wszystkie wyniki badań spektroskopowych dotyczące wodoru i jonów wodoropodobnych. Tymczasem stosując spektrografy o dużej zdolności rozdzielczej (por. § 28.4) stwierdzono, że obserwowane linie widmowe nie są pojedyncze, lecz składają się z dwóch lub więcej linii o bardzo bliskich długościach fali. Temu rozszczepieniu linii widmowych nadano nazwę *subtelnej struktury widma.* Zaobserwowane zjawisko próbowano wyjaśnić wykorzystując model Bohra — niestety bez rezultatów. Wyjaśnienie nowych faktów doświadczalnych wymagało wprowadzenia nowych założeń i uzupełnień, czyli konieczne było przejście do obrazu modelowego bardziej skomplikowanego.

Z mechaniki wiadomo, że pod wpływem działania sił odwrotnie proporcjonalnych do kwadratu odległości (np. siły kulombowskie, grawitacyjne) mogą występować nie tylko ruchy po kole. Sommerfeld wprowadził do modelu atomu orbity eliptyczne. W jednym z ognisk elipsy, zakreślanej przez elektron, znajduje się jądro atomowe. Koło jest specjalnym przypadkiem elipsy, której oba ogniska zlały się w jednym punkcie. A zatem przypadek, uwzględniający kołowy ruch elektronu zbadany przez Bohra, jest szczególnym przypadkiem ujęcia ogólniejszego, podanego przez Sommerfelda.

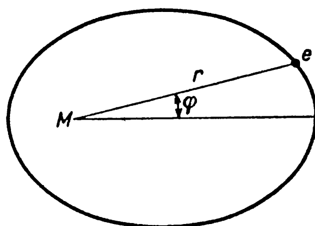
Rozważając ruch płaski elektronu po elipsie we współrzędnych biegunowych r i φ

(rys. 30.8) uwzględniać należy tzw. pęd radialny p_r, wynikający ze zmiany r przy zmianie φ (nie istniejący w przypadku ruchu kołowego) oraz moment pędu p_φ, wynikający z ruchu po orbicie. Sommerfeld zakłada, że obie wielkości p_r i p_φ są skwantowane. Innymi słowy, zastąpienie ruchu kołowego ruchem eliptycznym zwiększyło liczbę warunków kwantowych do dwóch. Oba warunki kwantowe można formułować w różny sposób. Wygodnie nam będzie podać je od razu w postaci zgodnej z wynikami mechaniki falowej. Jedna z liczb kwantowych będzie się wiązała z energią elektronu na orbicie, druga — z kształtem orbity i z momentem pędu na orbicie (por. wzór (30.15)). Pierwszą z nich w dalszym ciągu będziemy oznaczać literą n i nazywać *liczbą kwantową główną*:

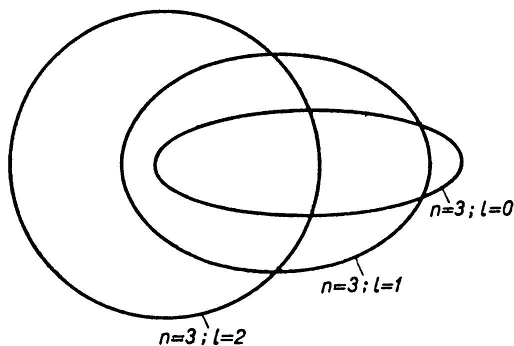
$$n = 1, 2, 3, \ldots \infty.$$

Drugą liczbę nazywać będziemy *liczbą kwantową poboczną lub azymutalną* i oznaczać symbolem l. Zakładamy, że liczba l przyjmuje wartości

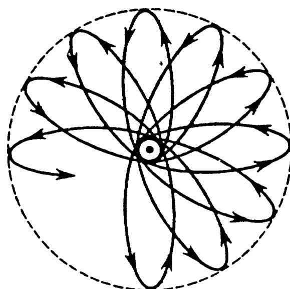
$$l = 0, 1, 2, \ldots (n-1).$$



Rys. 30.8



Rys. 30.9



Rys. 30.10

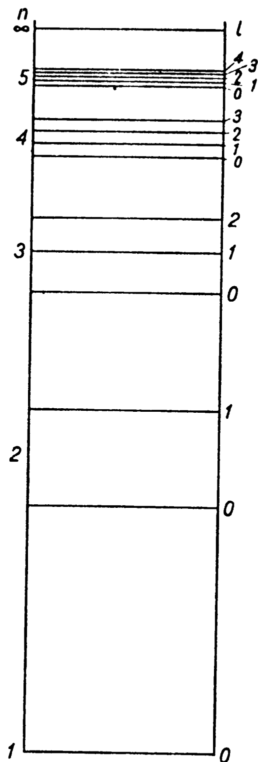
Liczba kwantowa główna określa możliwe wartości energii całkowitej elektronu na elipsie. Pierwotnie zakładano, że energia E_n na orbicie eliptycznej wyraża się wzorem

$$E_n = -\frac{me^4Z^2}{8\pi\varepsilon_0^2n^2h^2}, \tag{30.14}$$

a więc wzorem całkowicie zgodnym z wzorem (30.9), wyrażającym energię elektronu na orbicie kołowej w atomie wodoru. (Wprowadzenie czynnika Z^2 uogólnia wzór (30.9) na pierwiastki o liczbie atomowej $Z > 1$.) Założenie to jest uzasadnione, gdyż z praw

dynamiki wynika, że energia ta zależy tylko od długości wielkiej osi elipsy, a nie zależy od jej kształtu.

Liczba kwantowa l określa małą oś elipsy. Im ta liczba jest mniejsza, tym elipsa jest bardziej wydłużona; gdy $l = n-1$, elipsa przechodzi w koło. Na przykład dla $n = 3$ mamy trzy możliwe wartości l, a mianowicie: 0, 1 i 2. Pierwszym dwóm wartościom odpowiadają elipsy różnych kształtów (rys. 30.9), trzeciej — koło.



Rys. 30.11

Początkowo moment pędu elektronu na orbicie wyrażano następującym wzorem:

$$p = l\frac{h}{2\pi}. \tag{30.15}$$

Innymi słowy, w tym ujęciu *liczba l przedstawia moment pędu na orbicie wyrażony w jednostkach* $h/2\pi$. Z tego uproszczonego ujęcia będziemy dalej korzystali (por. omówienie rys. 30.13), aczkolwiek w ujęciu mechaniki kwantowej całkowity moment pędu spełnia równanie

$$p = \frac{h}{2\pi}\sqrt{l(l+1)}. \tag{30.16}$$

Dalsza konfrontacja wyników teoretycznych i doświadczalnych wykazała, że zachodzi konieczność wniesienia nowej poprawki do stworzonego obrazu. Prędkość elektronu na orbitach osiąga duże wartości. Na przykład w przypadku elektronu krążącego na podstawowym torze w atomie wodoru stosunek jego prędkości v do prędkości światła c wynosi, zgodnie z wzorem (30.7),

$$\frac{v}{c} = \frac{e^2}{2nh\varepsilon_0 c} = \frac{1}{137},$$

czyli $v > 2000$ km/s. Wobec tak dużych prędkości zachodzi konieczność relatywistycznego uwzględniania zmian masy. Sommerfeld uwzględnił tę poprawkę i wykazał, że ruch elektronu odbywa się po torze przedstawionym na rys. 30.10.

W tych warunkach energia elektronu na torze zależy już od kształtu elipsy. Widać to ze wzoru Sommerfelda na energię elektronu na orbicie n w atomie wodoru:

$$E_n = -\frac{me^4}{8\pi\varepsilon_0^2 n^2 h^2}\left[1 + \frac{1}{137n}\left(\frac{n}{l+1} - \frac{3}{4}\right)\right]. \tag{30.17}$$

We wzorze tym występują obie liczby kwantowe n i l. Warto dłużej zatrzymać się przy tej zależności. Ze wzoru (30.17) wynika, że różnym wartościom l przy tym samym n odpowiadają różne energie, czyli że poziom energetyczny odpowiadający liczbie n, który dotychczas uważaliśmy za pojedynczy, ulega rozszczepieniu na kilka bliskich siebie poziomów. Poziom $n = 1$ nie ulega rozszczepieniu, gdyż odpowiada mu tylko jedna wartość $l = 0$. Rozszczepienie poziomów energetycznych przedstawia rys. 30.11.

Tłumaczy ono rozszczepienie linii widmowych, czyli strukturę subtelną widma, o której już wcześniej była mowa.

Przejście od modelu Bohra do modelu bardziej skomplikowanego z poprawkami Sommerfelda rozszerzyło więc zakres zjawisk, które można tłumaczyć modelowo. Bohr wprowadzając jedną liczbę kwantową wyjaśnił mechanizm powstawania serii widmowych w wodorze; Sommerfeld wprowadzając dwie liczby kwantowe był już w stanie wytłumaczyć subtelną strukturę tego widma.

Jednak dalsze jeszcze fakty doświadczalne wymagały wyjaśnienia z punktu widzenia modelowej budowy atomu, m.in. zjawisko Zeemana. Do tych spraw wrócimy jeszcze w § 30.9.

30.7. Fale materii

Zanim przejdziemy do dalszego rozszerzania wiadomości dotyczących elektronu w atomie powrócimy jeszcze do sensu fizycznego, jaki kryje się poza matematyczną zależnością wyrażającą pierwszy postulat Bohra, a mianowicie

$$rmv = n\frac{h}{2\pi}.$$

Rozważania teoretyczne Ludwika de Broglie'a (1924) oraz badania doświadczalne zapoczątkowane przez Davissona i Germera doprowadziły do tego, że ten intuicyjnie przez Bohra odgadnięty warunek nabrał konkretnego znaczenia.

Przypominamy, że podstawowe doświadczenia dotyczące zachowania się promieni katodowych w polu elektrycznym i magnetycznym wykazały, że elektrony zachowują się jak naładowane ujemnie cząstki o pewnej masie, podlegające zarówno prawom elektrodynamiki, jak i prawom dynamiki Newtona.

Odmienny charakter wykazują elektrony w innych doświadczeniach. Davisson i Germer badali rozpraszanie wiązek elektronowych na powierzchni kryształków niklu. Stwierdzili oni nierównomierny rozkład przestrzenny rozproszonych elektronów. Występowało „selektywne odbicie", analogiczne do obserwowanego w metodzie Braggów dla promieni Röntgena. Powtarzając doświadczenie w zmienionym układzie, a mianowicie przepuszczając wiązkę elektronów przez cienkie folie złota (struktura mikrokrystaliczna) otrzymano zarówno dla szybkich (Thomson), jak i dla powolnych (Tartakowski) elektronów obrazy dyfrakcyjne w postaci pierścieni, przypominających pierścienie uzyskiwane w metodzie Debye'a i Scherrera dla promieni Röntgena. W doświadczeniach powyższych wystąpiły charakterystyczne zjawiska falowe: ugięcie i interferencja. Trzeba więc było przypisać elektronowi naturę falową, niezależnie od wspomnianej poprzednio natury korpuskularnej.

Ludwik de Broglie już wcześniej, gdyż w roku 1924, wysunął hipotezę, że każdej poruszającej się cząstce można przypisać falę pewnej długości, przy czym długość fali i jej częstotliwość związane są z pędem i energią cząstki zależnościami analogicznymi do tych, które obowiązują dla fotonów, a mianowicie:

$$E_{\text{fotonu}} = h\nu; \qquad \text{pęd fotonu} = mc = \frac{h\nu c}{c^2} = \frac{h\nu}{c} = \frac{h}{\lambda}. \tag{30.18}$$

Przekształcając wzór na pęd opieramy się na zależności między masą i energią, wynikającej z teorii względności:

$$E = mc^2$$

(por. § 7.6).

Pęd p cząstki o masie m i prędkości v wynosi zatem

$$p = mv = \frac{h}{\lambda}, \tag{30.19}$$

skąd

$$\lambda = \frac{h}{mv} = \frac{h}{p}. \tag{30.20}$$

Wspomniane wyżej doświadczenia Davissona i Germera oraz Thomsona i Tartakowskiego są pięknym doświadczalnym potwierdzeniem hipotezy de Broglie'a. Załóżmy na przykład, że elektron o początkowej prędkości $v = 0$ przelatuje przez pole elektryczne o różnicy potencjałów U. Wtedy

$$\tfrac{1}{2}mv^2 = eU,$$

skąd

$$v = \sqrt{\frac{2eU}{m}}, \tag{30.21}$$

a długość fali zgodnie z wzorem (30.20)

$$\lambda = \frac{h}{\sqrt{2eUm}}. \tag{30.22}$$

Po podstawieniu danych liczbowych otrzymujemy

$$\lambda = \sqrt{\frac{150}{U}}\ \text{Å},$$

gdy U jest wyrażone w woltach. Na przykład dla napięcia $U = 150$ V otrzymujemy $\lambda = 1$ Å, a więc długość fali elektronu w tych warunkach jest tego samego rzędu co promieni Röntgena.

Elektron wykazuje zatem podwójną naturę — w pewnych doświadczeniach zachowuje się jak cząstka materialna, w innych — jak fala. Można więc mówić o *dualizmie korpuskularno-falowym* elektronu, a nawet rozszerzać to stwierdzenie na inne cząstki, jak np. protony, atomy, cząsteczki (doświadczenia Sterna, Johnsona). Przypominamy, że już poprzednio podkreślaliśmy dualizm natury światła. Światło w pewnych okolicznościach (np. zjawisko fotoelektryczne) wykazuje cechy korpuskularne — fotonom przypisujemy masę, energię i pęd. W innych warunkach światło ma cechy fali elektromagnetycznej (zjawisko ugięcia, interferencji, polaryzacji).

Zastosujmy uzyskany obraz falowy elektronu do modelu Bohra. Elektron porusza się po orbicie kołowej o promieniu r z prędkością v. Według de Broglie'a przypisujemy mu długość fali $\lambda = h/mv$. Tylko wtedy fale elektronowe związane z obiegiem elektronu po orbicie $2\pi r$ nie będą się wygaszały podczas interferencji, gdy długość obwodu $2\pi r$ będzie całkowitą wielokrotnością długości fali. Przedstawiają to poglądowo rys. 30.12a,

b i c. Dla fal przedstawionych na rysunkach a i b spełniony jest warunek całkowitej wielokrotności — fale się nie wygaszają, na rysunku c warunek ten nie jest spełniony — fale nakładają się we wszelkich możliwych fazach i ostatecznie ulegają wygaszeniu.

Warunek kwantowy dla dozwolonych orbit elektronowych, wynikający z teorii de Broglie'a, ma postać

$$2\pi r = n\lambda, \quad \text{gdzie} \quad n = 1, 2, 3, \ldots \infty. \tag{30.23}$$

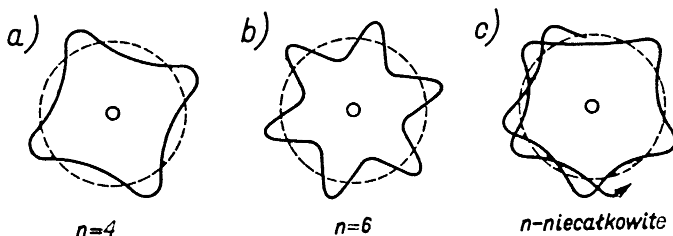
Można go doprowadzić do postaci podanej przez Bohra, gdyż na podstawie wzoru (30.20)

$$2\pi r = n\frac{h}{mv},$$

skąd

$$rmv = n\frac{h}{2\pi}.$$

Jest to znana postać matematyczna pierwszego postulatu Bohra (moment pędu elektronu na orbicie stacjonarnej jest całkowitą wielokrotnością $h/2\pi$).



Rys. 30.12

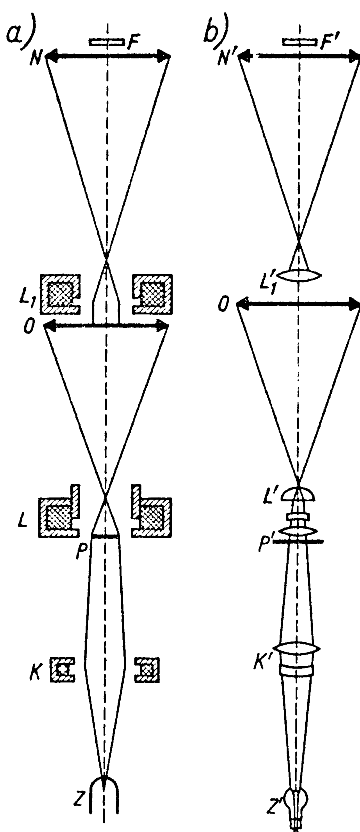
30.7.1. *Mikroskop elektronowy*. Omawiając w punkcie 25.7.1 zagadnienie zdolności rozdzielczej mikroskopu optycznego stwierdziliśmy, że zależy ona od długości fali zastosowanego promieniowania i w przypadku obserwacji fotograficznej w świetle nadfioletowym pozwala na rozróżnianie punktów przedmiotu odległych od siebie o około 100 nm.

Stwierdzenie dualizmu korpuskularno-falowego elektronu otworzyło nowe możliwości przed badaniami mikroskopowymi. Stosując napięcie przyspieszające elektrony, rzędu dziesiątków kilowoltów, uzyskać można takie prędkości ruchu elektronów, którym odpowiadają długości fal rzędu 10^{-11} m (czyli dziesiątych części Å). Zastosowanie takich fal w mikroskopii znacznie zwiększyłoby oczywiście zdolność rozdzielczą odpowiednich mikroskopów.

Zasada mikroskopu elektronowego sprowadza się właśnie do tego, że wiązce elektronów uzyskanej dzięki termoemisji (por. § 32.6) nadaje się energię rzędu dziesiątków kiloelektronowoltów i skierowuje się ją na niesłychanie cienką warstwę badanego ciała. Dzięki temu elektrony przechodzą przez tę warstwę z małą stratą energii. Podczas tego przenikania przez materię elektrony ulegają ugięciu. Ugięta wiązka elektronowa przechodzi dalej przez odpowiednio dobrane co do kształtu i rozkładu natężeń obszary pól magnetycznych i elektrycznych, stanowiących tzw. *soczewki elektronowe*. Soczewki te działają podobnie jak soczewka obiektywowa i okularowa w mikroskopie optycznym, tzn.

wytwarzają powiększony „elektronowy" obraz przedmiotu na fluoryzującym ekranie lub płycie fotograficznej ustawionej na miejscu ekranu. Rysunek 30.13 przedstawia schemat mikroskopu elektronowego i mikroskopu optycznego. W obu mikroskopach znajdujemy te same elementy, jak to widać z następującego zestawienia:

Z i Z' — źródło elektronów i źródło światła,
P i P' — przedmioty podlegające obserwacji,
L i L' — obiektyw magnetyczny i obiektyw optyczny,
O i O' — obrazy pośrednie wytworzone w obu mikroskopach,
L_1 i L_1' — soczewki projekcyjne obu urządzeń,
N i N' — obrazy końcowe wytworzone na płytach fotograficznych lub na płycie F' i ekranie fluoryzującym F.



Rys. 30.13

Rozmiary mikroskopu elektronowego są znacznie większe od rozmiarów mikroskopu optycznego. Wysokość przyrządu jest rzędu 1 m. Wewnątrz przyrządu utrzymuje się wysoką próżnię w celu uniknięcia zderzeń elektronów z cząsteczkami powietrza. Rząd wielkości próżni jest 10^{-3} N/m^2 (czyli około 10^{-5} mm Hg). Stosowane napięcia sięgają 50 kV.

Zastosowania mikroskopu elektronowego w różnego rodzaju badaniach naukowych, w biologii, w medycynie itp. stale się rozszerzają. Opracowywane są coraz to nowe metody obserwacji mikroskopowej. W odniesieniu do cienkich obiektów o grubościach ułamków mikrometrów można stosować prześwietlanie preparatu wiązkami elektronowymi. Badany obiekt jest wtedy umieszczany na odpowiednio wytwarzanej „podkładce", o grubości rzędu setnych części mikrometra.

Aby uzyskać większą kontrastowość otrzymywanych obrazów, stosuje się czasem ukośne napylanie obiektu w próżni różnymi metalami. W ten sposób uzyskuje się „wycieniowanie" danego obiektu warstwą osadzonego metalu, który następnie silnie rozpraszając padające elektrony podnosi kontrastowość obrazu.

Do badania powierzchni ciał stałych stosuje się specjalną metodę, tzw. repliki, polegającą na nanoszeniu delikatnej warstwy (z różnych metali, kwarcu, związków organicznych) na badaną powierzchnię, a następnie zdjęciu tej warstwy z zachowaniem wielkiej ostrożności i poddaniu jej badaniom mikroskopowym.

Mikroskop elektronowy pozwala już obecnie na „oglądanie" dużych cząsteczek w postaci oddzielnych plamek na ekranie. Stosując dodatkowe mikroprojektory można nawet badać kształt makrocząsteczek. W tych warunkach można już rozróżniać punkty leżące w odległościach 2–3 Å.

30.8. Moment magnetyczny elektronu

Model Bohra budowy atomu można wykorzystać do poglądowego wprowadzenia wielkości odgrywającej podstawową rolę w mechanice kwantowej, a mianowicie — orbitalnego momentu magnetycznego elektronu, zwanego czasem krótko *momentem magnetycznym*. Elektron poruszający się po zamkniętej orbicie dokoła jądra można rozpatrywać jako prąd elektryczny o pewnym natężeniu I, którego obwód obejmuje powierzchnię S. Z rozważań § 21.6 wynika, że z takim prądem można kojarzyć istnienie momentu magnetycznego $\mu = IS$. Wektor $\boldsymbol{\mu}$ skierowany jest prostopadle do płaszczyzny obwodu.

Na podstawie założeń Bohra można obliczyć, że moment magnetyczny μ_e elektronu jest całkowitą wielokrotnością pewnej wielkości, której nadano nazwę *magnetonu Bohra* μ_B, a mianowicie

$$\mu_e = l\mu_B, \tag{30.24}$$

gdzie l oznacza liczbę kwantową poboczną (charakteryzującą orbitę, po której elektron odbywa ruch, i moment pędu na tej orbicie). Magneton Bohra μ_B równa się

$$\mu_B = \frac{e}{2m}\frac{h}{2\pi}, \tag{30.25}$$

gdzie e i m oznaczają odpowiednio ładunek i masę spoczynkową elektronu. Wartość liczbowa magnetonu Bohra wynosi $9{,}273 \cdot 10^{-24}$ J/T*. Wzór (30.24) został nieco zmie-

* Magneton Bohra jest też często wyrażany w $\mathrm{A \cdot m^2}$. Łatwo wykazać, że obie jednostki są równoważne ($\mathrm{A \cdot m^2 \equiv J/T}$).

niony w ujęciu mechaniki kwantowej, a mianowicie

$$\mu_e = \mu_B \sqrt{l(l+1)}\,. \tag{30.26}$$

Warto od razu podkreślić, że elektron prócz orbitalnego momentu magnetycznego ma jeszcze moment magnetyczny spinowy, związany z ruchem obrotowym elektronu dokoła własnej osi.

30.9. Liczba kwantowa magnetyczna

W § 30.7 wspomnieliśmy, że wprowadzone poprzednio dwa warunki kwantowe, wynikające z uwzględnienia płaskiego ruchu elektronu po orbicie eliptycznej, nie wystarczają np. do wyjaśnienia zjawiska Zeemana. Wrócimy obecnie do tego zjawiska.

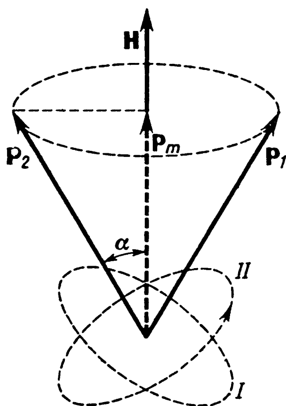
W roku 1896 Zeeman stwierdził, że linie widmowe ulegają rozszczepieniu, gdy ciało wysyłające promieniowanie poddane jest działaniu silnego pola magnetycznego. Zjawisko to było częściowo wyjaśnione na gruncie elektrodynamiki klasycznej. Sommerfeld dał wyjaśnienie kwantowe tzw. normalnego zjawiska Zeemana, wprowadzając trzecią liczbę kwantową, tzw. *liczbę kwantową magnetyczną*, związaną z kwantowaniem przestrzennym. Spróbujemy krótko przedstawić bieg rozumowania.

W dotychczasowych rozważaniach uwzględnialiśmy tylko ruch płaski elektronu po kole lub elipsie (czy też po rozecie eliptycznej — rys. 30.10) odbywający się w stałej płaszczyźnie, bez określenia położenia tej płaszczyzny w przestrzeni. Z punktu widzenia mechaniki klasycznej wszystkie ustawienia płaszczyzny toru elektronu w przestrzeni są równie możliwe. Sommerfeld przewidział istnienie tylko pewnych dozwolonych — skwantowanych położeń. Żeby jednak mówić o określonych, dozwolonych położeniach danej płaszczyzny w przestrzeni, trzeba mieć jakiś układ odniesienia, trzeba ustalić jakiś wybrany kierunek odniesienia. Ten wyróżniony kierunek występuje wyraźnie wtedy, gdy atom znajduje się w polu magnetycznym. Pole takie zakłóca ruch elektronu, ale wywołane zakłócenie można zaniedbać, gdy pole jest bardzo słabe. Zakłócający wpływ pola magnetycznego wiąże się ściśle z momentem magnetycznym orbitalnym, jaki w poprzednim punkcie przypisaliśmy elektronowi. Moment magnetyczny elektronu można myślowo kojarzyć z małym dipolem magnetycznym, umieszczonym w środku orbity prostopadle do jej powierzchni. Na dipol magnetyczny działa w polu magnetycznym para sił usiłująca obrócić go tak, aby ustawił się w kierunku pola **H**. Trzeba jednak pamiętać, że omawiany dipol wynika z krążenia elektronu po orbicie. Z ruchem tym wiąże się moment pędu również prostopadły do płaszczyzny orbity. W wyniku działania pary sił na dipol (związany z poruszającym się po orbicie elektronem) rozpoczyna się precesja, podczas której wektor momentu magnetycznego i wektor momentu pędu (oba prostopadłe do płaszczyzny orbity) opisują pobocznicę stożka dokoła kierunku pola magnetycznego jako osi (rys. 30.14). Każdemu ustawieniu w przestrzeni wektora momentu pędu odpowiada prostopadłe ustawienie płaszczyzny orbity elektronu. A więc podczas precesji odbywa się w przestrzeni ciągła zmiana położenia płaszczyzny orbity, z zachowaniem jej prostopadłości do wektora momentu pędu **p**, który jest stale odchylony od kierunku **H** o kąt α. Takie dwa położenia płaszczyzny orbity *I* i *II* i związane z nimi położenia wektorów $\mathbf{p}_1$ i $\mathbf{p}_2$ zaznaczone są na rys. 30.14.

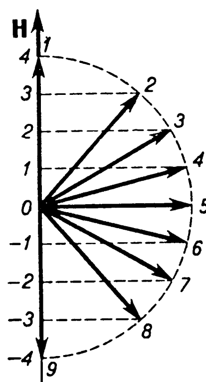
Istnienie pola magnetycznego wpływa na stan energetyczny elektronu. Zmiana energii elektronu wywołana istnieniem pola magnetycznego zależy od ustawienia płaszczyzny orbity elektronu względem kierunku pola, co można ująć inaczej, mówiąc o zależności ΔE od kąta α, czyli od kąta między kierunkiem pola **H** i momentem magnetycznym elektronu (lub jego momentem pędu). Zmiany energii zachodziłyby w sposób ciągły, gdyby wszystkie ustawienia płaszczyzny orbity w przestrzeni (wszystkie wartości kąta α) były możliwe. Wystąpiłoby wtedy poszerzenie linii widmowych, rozciągnięcie ich w rodzaj pasma. Wyniki badań doświadczalnych są inne: w polu magnetycznym zamiast rozszerzenia linii obserwuje się ich rozszczepienie. Aby wyjaśnić to zjawisko, trzeba wprowadzić założenie o *kwantowaniu przestrzennym*, a mianowicie, aby być w zgodzie z doświadczeniem, trzeba przyjąć, że tylko takie ustawienia orbitalnych momentów pędów w przestrzeni są możliwe, przy których ich rzut p_m na wyróżniony kierunek pola magnetycznego jest całkowitą wielokrotnością wielkości $h/2\pi$, czyli

$$p_m = m\frac{h}{2\pi}, \tag{30.27}$$

gdzie m oznacza liczbę kwantową magnetyczną.



Rys. 30.14



Rys. 30.15

Z rysunku 30.14 widać, że rzut p_m momentu pędu **p** na kierunek pola **H** równa się

$$p_m = p\cos\alpha,$$

gdzie α jest kątem między kierunkiem pola i kierunkiem wektora **p**. Kąt α może przyjmować skwantowane wartości w granicach od zera do π, a zatem $\cos\alpha$* ma kwantowe wartości zarówno dodatnie, jak i ujemne. Liczba kwantowa magnetyczna m może przyjmować wartości całkowite dodatnie lub ujemne, w granicach od $-l$ do $+l$, tzn.

* Warunek kwantowy, jaki zgodnie z mechaniką kwantową musi spełniać $\cos\alpha$, ma następującą postać matematyczną:

$$\cos\alpha = \frac{m}{\sqrt{l(l+1)}}.$$

ogółem $(2l+1)$ wartości

$$-l, -l+1, -l+2, \ldots, 0, 1, \ldots, l-1, l.$$

Przykład kwantowania przestrzennego dla liczby kwantowej $l = 4$ przedstawiony jest na rys. 30.15. Liczba kwantowa $l = 4$ oznacza, że orbitalny moment pędu elektronu wynosi $4h/2\pi$. Przyjmując zatem za jednostkę wielkość $h/2\pi$ odłożymy moment pędu jako wektor o długości równej czterem jednostkom. Gdyby nie było ograniczeń związanych z kwantowaniem przestrzennym, wektor ten, o początku w punkcie O, mógłby się ustawiać dowolnie, a zatem jego koniec przypadałby na kuli o promieniu równym 4 jednostkom. W rzucie na płaszczyznę rysunku koniec wektora **p** mógłby zajmować dowolne położenia na kole (o promieniu r równym 4 jednostkom) zakreślonym z punktu O. Na rys. 30.15 zaznaczone jest tylko półkole.

Kwantowanie przestrzenne sprowadza się do: 1) uwzględnienia wyróżnionego kierunku (kierunek **H** na rysunku) i 2) wybrania spośród wszystkich ustawień przestrzennych wektora **p** tylko takich, przy których rzut jego na kierunek **H** równa się całkowitej liczbie (jednostek $h/2\pi$), a więc 4, 3, 2, 1, 0, −1, −2, −3, −4. Widzimy, że dla $l = 4$ mamy 9 takich możliwych skwantowanych położeń.

30.10. Liczba kwantowa spinowa

W związku z powiększeniem zdolności rozdzielczej aparatur spektroskopowych ujawniły się nowe fakty związane ze strukturą widm, których nie można było wyjaśnić stosując omówione wyżej trzy liczby kwantowe. W roku 1925 Uhlenbeck i Goudsmit zaproponowali, aby elektron traktować jako naładowaną cząstkę wykonującą obrót dokoła własnej osi (model przybliżony — kulka o promieniu rzędu 10^{-15} m, wirująca dokoła własnej osi). Z obrotem elektronu dokoła własnej osi wiąże się jego własny, nieorbitalny, moment pędu, zwany krótko *spinem elektronu*. Spin jest również skwantowany: może on przyjmować wartości $sh/2\pi$, gdzie s oznacza liczbę kwantową spinową. Zgodnie z mechaniką kwantową i w celu otrzymania wyników zgodnych z doświadczeniami należało liczbie spinowej s przypisać wartość $+\frac{1}{2}$ lub $-\frac{1}{2}$. Często w skrócie mówi się po prostu, że spin elektronu jest połówkowy. Doświadczenia wykazują, że własny moment pędu elektronu (czyli spin) w polu magnetycznym może się ustawiać tylko równolegle lub antyrównolegle względem kierunku pola, z czym wiąże się występowanie dwóch wartości liczby kwantowej spinowej $+\frac{1}{2}$ i $-\frac{1}{2}$. W atomie zawierającym wiele elektronów możliwe są tylko równoległe lub antyrównoległe ustawienia ich spinów.

Po tym ogólnym przeglądzie warunków kwantowych sprawdźmy, jak dalece skomplikował się obraz poziomów energetycznych atomu, gdy odeszliśmy od pierwotnej koncepcji Bohra.

Weźmy dla przykładu stan energetyczny elektronu określony liczbą kwantową $n = 3$. Według Bohra odpowiada mu pojedynczy poziom energetyczny związany z jedną ściśle określoną wartością energii. Biorąc pod uwagę poprawki Sommerfelda trzeba uwzględnić możliwość wystąpienia przy danym n trzech różnych wartości l:

0, 1 i 2. Z przyporządkowania poszczególnym wartościom l różnych liczb kwantowych magnetycznych, a mianowicie:

$$l = 0 \quad m = 0,$$
$$l = 1 \quad m = -1, 0, +1,$$
$$l = 2 \quad m = -2, -1, 0, +1, +2,$$

wyniknie łącznie 9 możliwych poziomów energetycznych. Ale z każdym z 9 wymienionych poziomów energetycznych można związać 2 różne ustawienia spinów elektronowych. Ogółem zatem zamiast jednego poziomu energetycznego, wynikającego z teorii Bohra, mamy 18 możliwych poziomów (§ 30.11).

Na zakończenie zestawmy raz jeszcze mocne i słabe punkty teorii Bohra–Sommerfelda. Daje ona proste wyjaśnienie mechanizmu powstawania serii widmowych wodoru, tłumaczy subtelną strukturę widma wodoru, normalne zjawisko Zeemana, analogiczne do niego zjawisko Starka, polegające na rozszczepieniu linii widmowych przy zastosowaniu pola elektrycznego, daje wyniki przy analizie widm atomów wodoropodobnych. Osiągnięcia te są możliwe dzięki wprowadzaniu coraz to nowych „kwantowych" założeń, dobieranych tak, aby były potwierdzone doświadczalnie. W zastosowaniu do tłumaczenia bardziej skomplikowanych zjawisk omawiana teoria zawodzi. Słabym jej punktem jest niewątpliwie to, że opiera się ona równocześnie na 2 sprzecznych podstawach: na mechanice klasycznej, rządzącej ruchem elektronów na orbicie, i na teorii kwantowej, z której wynika istnienie orbit stacjonarnych i możliwość obliczania częstotliwości wysyłanego promieniowania. Warto jednak podkreślić, że teoria Bohra spełniła bardzo poważne zadanie: wpłynęła pobudzająco na rozwój badań w dziedzinie fizyki atomowej i przyczyniła się do ugruntowania kwantowego punktu widzenia w fizyce współczesnej.

Po roku 1925 teoria Bohra ustąpiła miejsca nowym teoriom, tzw. mechanice kwantowej Heisenberga oraz mechanice falowej Schrödingera. Są to jednak teorie bardziej matematyczne, bardziej abstrakcyjne i dlatego jeszcze dziś w zagadnieniach, w których możliwe jest poglądowe przedstawienie zagadnienia, powołujemy się na obraz modelowy Bohra.

30.11. Zasada Pauliego. Układ okresowy pierwiastków

Jak wynika z poprzednich paragrafów, elektrony w atomie spełniają 4 warunki kwantowe, związane z czterema liczbami kwantowymi. Są to:

n — liczba kwantowa główna,

$$n = 1, 2, 3 \ldots \infty,$$

związana z energią na orbicie;

l — liczba kwantowa poboczna, zwana też azymutalną lub orbitalną,

$$l = 0, 1, 2, \ldots, n-1,$$

związana z momentem pędu na orbicie i kształtem elipsy;

m — liczba kwantowa magnetyczna,

$$m = -l, (-l+1), \dots 0, \dots, +l,$$

związana z rzutem momentu pędu na orbicie na kierunek pola magnetycznego;

s — liczba kwantowa spinowa,

$$s = +\tfrac{1}{2}, -\tfrac{1}{2},$$

związana z rzutem spinu elektronu na kierunek pola.

Na podstawie bogatego materiału doświadczalnego Pauli w r. 1925 sformułował swój słynny „zakaz", zwany *zasadą Pauliego.* Zgodnie z tą zasadą *w atomie nie mogą istnieć dwa elektrony o jednakowych wszystkich czterech liczbach kwantowych.*

Stosując zasadę Pauliego można przewidzieć rozbudowę kolejnych powłok elektronowych w atomach pierwiastków. Pojęciem *powłoki elektronowej* obejmujemy zespół poziomów energetycznych odpowiadających liczbie kwantowej głównej *n*. Często też są w użyciu następujące symboliczne oznaczenia literowe:

liczba kwantowa główna *n*	1	2	3	4
symbol powłoki	*K*	*L*	*M*	*N* itd.

W powłoce elektronowej *K* mogą występować tylko elektrony o $l = 0$ i $m = 0$. Spinowa liczba kwantowa *s* przyjmuje dwie wartości: $+\frac{1}{2}$ i $-\frac{1}{2}$. A zatem w powłoce *K* mogą się znaleźć najwyżej dwa elektrony. Atom helu ($Z = 2$) ma właśnie dwa elektrony wypełniające całkowicie powłokę *K*.

W drugiej z kolei powłoce *L* są już możliwe dwie wartości l: $l = 0$ i $l = 1$. Każdej wartości *l* przypisujemy zespół poziomów energetycznych stanowiący podpowłokę. Powłoka *L* składa się więc z dwóch podpowłok. Wprowadzamy również symbolikę literową dla rozróżnienia podpowłok, a mianowicie:

liczba kwantowa poboczna *l*	0	1	2	3	4	5
symbol podpowłoki	*s*	*p*	*d*	*f*	*g*	*h* itd.

Podpowłoce *s* odpowiada jedna tylko liczba $m = 0$, ale dwie liczby spinowe: $+\frac{1}{2}$ i $-\frac{1}{2}$, czyli w tej podpowłoce znajdują się znów tylko dwa elektrony różniące się liczbami spinowymi. Elektrony te oznaczamy symbolicznie jako elektrony 2*s*: liczba oznacza numer kolejny powłoki odpowiadający wartości *n*, litera określa podpowłokę. Rozbudowa elektronowa podpowłoki *s* w powłoce 2, czyli w powłoce *L*, odbywa się w licie ($Z = 3$) i berylu ($Z = 4$).

W podpowłoce drugiej, czyli w podpowłoce *p* ($l = 1$) mamy już trzy możliwe wartości *m*: -1, 0 i $+1$. Poza tym każdy z tych elektronów ma jeszcze dwie możliwości ustawienia spinów, a zatem ogółem w tej podpowłoce znajduje się 6 elektronów. Rozbudowa drugiej podpowłoki w powłoce *L* przypada na bor, węgiel, azot, tlen, fluor i neon. Neon, dziesiąty pierwiastek układu okresowego, ma ogółem 10 elektronów w stanie normalnym atomu, z tego 2 o symbolach 1*s* (w powłoce *K*), 2 — o symbolach 2*s* (w podpowłoce *s* powłoki *L*) i 6 — o symbolach 2*p* (w podpowłoce *p* powłoki *L*).

Analogiczne rozumowania można przeprowadzić dla dalszych powłok *M*, *N*, *O* itd.

Łatwo ustalić maksymalną liczbę elektronów w podpowłokach i powłokach. W powłoce scharakteryzowanej liczbą kwantową *n* mamy tyle podpowłok, ile jest różnych wartości *l*, czyli znowu *n* podpowłok. W podpowłoce o liczbie kwantowej pobocznej *l*

mamy $(2l+1)$ możliwych wartości m i dla każdego m jeszcze możliwe dwie liczby spinowe, czyli ogółem liczba elektronów w tej podpowłoce wynosi

$$2(2l+1).$$

Ogółem zaś w całej powłoce n będzie N elektronów:

$$N=\sum_{l=0}^{n-1}(2l+1)2=2\sum_{l=0}^{n-1}[2l+1]=2[1+3+5+\ldots+(2n-1)].$$

Uwzględniając wzór na sumę postępu arytmetycznego otrzymujemy

$$N=2\,\frac{1+2n-1}{2}\,n=2n^2. \tag{30.28}$$

Z wzoru tego wynika, że maksymalna liczba elektronów w poszczególnych powłokach wynosi:

symbol powłoki	K	L	M	N	O
maksymalna liczba elektronów	2	8	18	32	50

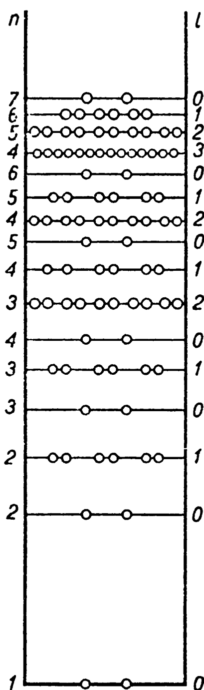
Tabela 30.1

Rozkład elektronów w atomach

Wartości n	1	2		3			4			
Wartości l	0	0	1	0	1	2	0	1	2	3
Symbol powłoki	K	L		M			N			
Symbol podpowłoki	1s	2s	2p	3s	3p	3d	4s	4p	4d	4f
Z Symbol pierwiastka										
1 H	1									
2 He	2									
3 Li	2	1								
4 Be	2	2								
5 B	2	2	1							
6 C	2	2	2							
7 N	2	2	3							
8 O	2	2	4							
9 F	2	2	5							
10 Ne	2	2	6							
11 Na	2	2	6	1						
12 Mg	2	2	6	2						
13 Al	2	2	6	2	1					
14 Si	2	2	6	2	2					
15 P	2	2	6	2	3					
16 S	2	2	6	2	4					
17 Cl	2	2	6	2	5					
18 A	2	2	6	2	6					
19 K	2	2	6	2	6		1			
20 Ca	2	2	6	2	6		2			
21 Sc	2	2	6	2	6	1	2			
22 Ti	2	2	6	2	6	2	2			

W tabeli 30.1 podane jest zestawienie kilkunastu pierwszych pierwiastków układu okresowego, z oznaczeniem elektronów wchodzących w skład ich atomów.

Dokładniejsza analiza tabeli pozwala zauważyć dość dziwny fakt w przypadku potasu ($Z = 19$). Gdybyśmy uważali za obowiązującą zasadę, że przechodzeniu do kolejnych pierwiastków o rosnącej liczbie atomowej Z towarzyszy stopniowa rozbudowa powłok i podpowłok o najmniejszych liczbach n i l, to można by się spodziewać, że po wypełnieniu w argonie ($Z = 18$) podpowłoki $l = 1$ w powłoce M rozpocznie się rozbudowa podpowłoki $l = 2$ w tej samej powłoce. Tymczasem (jak wynika z analizy widm) w potasie mimo pustej podpowłoki $l = 2$ pojawia się 19-ty elektron o charakterystyce $4s$, tzn. w powłoce N i podpowłoce $l = 0$. Podobna sytuacja rozbudowy powłok o większych liczbach n przed całkowitym wypełnieniem powłok poprzednich zachodzi w przypadku rubidu ($Z = 37$), srebra ($Z = 47$), cezu ($Z = 55$) i złota ($Z = 79$). Pozornie sprawia to wrażenie naruszenia pewnej regularności w rozbudowie atomów. W rzeczywistości sprawa sprowadza się do tego, że podstaw rozbudowy należy się dopatrywać w kolejnym wypełnianiu się poziomów energetycznych, począwszy od najniższych wartości energii. Z rozkładu poziomów energetycznych przedstawionych na rys. 30.11 wynikało, że wszystkie poziomy energetyczne związane z daną liczbą kwantową główną n są niższe od poziomów odpowiadających liczbie $n+1$. Na pewno jest to słuszne w przypadku atomu wodoru, ale nie obowiązuje dla wszystkich pierwiastków. W miarę rozbudowy atomu zmienia się rozkład poziomów energetycznych (rys. 30.16). Rysunek 30.16 przedstawia schematycznie rozkład energii w atomach z zaznaczeniem powłok (liczby n) i podpowłok (liczby l) oraz maksymalnej liczby elektronów w podpowłokach (małe kółeczka na liniach poziomych). Z rysunku widać, że poziom energetyczny $n = 4$ i $l = 0$ przypada niżej niż $n = 3$, $l = 2$, a więc 19-ty elektron potasu, elektron $4s$, zajmuje właściwe miejsce odpowiadające mniejszej energii.



Rys. 30.16

Podobne zakłócenia w rozkładzie poziomów energetycznych występują w pozostałych wymienionych wyżej przypadkach. Na przykład w rubidzie ostatni elektron jest elektronem $5s$, a nie spodziewanym $4d$ lub $4f$, co znów jest zgodne z rozkładem poziomów energetycznych.

Z punktu widzenia fizyki ważne jest rozróżnianie trzech stanów atomu: stanu normalnego, stanu wzbudzonego i stanu zjonizowania.

W stanie normalnym atomu wszystkie elektrony znajdują się na najniższych dozwolonych poziomach energetycznych.

W stanie wzbudzonym któryś z elektronów, dzięki uzyskaniu odpowiedniej porcji energii z zewnątrz, przeniesiony jest na poziom energetyczny wyższy. Po krótkim czasie atom wraca do stanu normalnego.

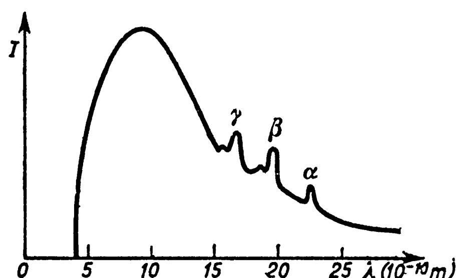
Gdy porcja energii dostarczona z zewnątrz jest wystarczająca do odrzucenia elektronu na poziom energetyczny zerowy, mówimy o stanie zjonizowania. Energie jonizacji zmieniają się okresowo, gdy przechodzimy od jednych pierwiastków układu okresowego do następnych. Pierwiastki o wybitnie trwałej budowie atomowej, a więc gazy szlachetne, odznaczają się dużymi energiami jonizacji. Stosunkowo małą energię jonizacji mają metale alkaliczne, łatwo odszczepiające swoje pojedyncze elektrony znajdujące się na ostatniej podpowłoce.

Podkreślmy na zakończenie tego punktu, że dotychczasowe rozważania dotyczyły izolowanego pojedynczego atomu. Elektronom w takim atomie przypisywaliśmy konkretne dozwolone poziomy energetyczne. Zasada Pauliego określała liczbę elektronów w atomie w poszczególnych powłokach i podpowłokach.

Jak zobaczymy (por. § 32.1), sprawa niesłychanie się komplikuje, gdy przechodzimy do gęstego zbioru atomów w ciele stałym.

30.12. Widma rentgenowskie

Zwracaliśmy już uwagę w punkcie 29.8.2, że promieniowanie rentgenowskie daje na spektrogramach *widmo ciągłe*, zwane *widmem hamowania*, w postaci słabo zaczernionego tła oraz *widmo liniowe*, tzw. charakterystyczne, zależne od materiału antykatody (por. rys. 29.23). Rysunek 30.17 przedstawia przykładowo rozkład natężeń w widmie



Rys. 30.17

promieniowania rentgenowskiego w zależności od długości fali. Maksima oznaczone literami α, β, γ odpowiadają liniom widma charakterystycznego. Ich powstawanie wyjaśnimy później. Drugą uderzającą cechą widma jest gwałtowne urywanie się widma ciągłego od strony fal krótkich. Wartości granicznej λ_{min} odpowiada maksymalna częstotliwość promieniowania ν_{max}, która, jak wykazały badania, spełnia zależność

$$\tfrac{1}{2}mv^2 = h\nu_{max}. \tag{30.29}$$

Lewa strona równania przedstawia maksymalną energię kinetyczną elektronu uzyskaną kosztem pracy sił pola w wyniku przelotu przez pole elektryczne o różnicy potencjałów U między antykatodą i katodą:

$$\tfrac{1}{2}mv^2 = eU. \tag{30.30}$$

Z porównania wzorów (30.29) i (30.30) wynika, że maksymalna energia kwantu promie-

niowania rentgenowskiego powstaje kosztem energii tych elektronów, które nie utraciły nic z zasobu uzyskanego kosztem pracy sił pola ani w drodze do antykatody, ani w czasie zderzeń nieelastycznych z atomami antykatody. Innymi słowy,

$$h\nu_{max} = eU,$$

$$\frac{hc}{\lambda_{min}} = eU,$$

skąd

$$\lambda_{min} = \frac{hc}{eU} \tag{30.31}$$

lub po podstawieniu danych liczbowych:

$$\lambda_{min} = \frac{123\,400}{U}\,\text{nm, gdy } U \text{ wyrażone jest w woltach.} \tag{30.32}$$

Ponieważ krótkofalowa granica widma jest ostro zarysowana, więc λ_{min} można dokładnie wyznaczać, a następnie z wzoru (30.31) obliczać np. stałą Plancka h. Wartości h otrzymane w ten sposób pokrywają się z wynikami otrzymanymi na innej drodze.

Po wyjaśnieniu pochodzenia krótkofalowej granicy widma rentgenowskiego przejdźmy do wyjaśnienia pochodzenia widma ciągłego. Elektrony dobiegające do antykatody tracą na ogół mniejszą lub większą część swej energii w sposób zupełnie dowolny, niekwantowy, podczas zderzeń nieelastycznych. Reszta energii elektronów w chwili ich zahamowania na powierzchni antykatody przekształca się w energię kwantów promieniowania rentgenowskiego i w związku z tym powstaje widmo ciągłe, zwane widmem hamowania.

Warto może raz jeszcze wrócić do równania (30.29). Odczytując je od lewej strony ku prawej powiemy, że kosztem energii kinetycznej padającego elektronu powstał kwant promieniowania rentgenowskiego; będzie to wyjaśnienie powstawania promieni Röntgena. Odczytując równanie od prawej strony ku lewej powiemy, że kosztem energii padającego kwantu promieniowania rentgenowskiego z atomu lub cząsteczki wyrzucony zostaje elektron o energii $\frac{1}{2}mv^2$. Tym razem opisujemy zjawisko fotoelektryczne w dziedzinie promieniowania rentgenowskiego. Przykładem takiego zjawiska jest jonizacja gazów pod wpływem promieni Röntgena.

Z kolei przechodzimy do widma rentgenowskiego charakterystycznego. W § 30.3 zapoznaliśmy się z modelem Bohra budowy atomu, jak również z wyjaśnieniem powstawania serii widmowych wodoru i wzorami określającymi długości fal poszczególnych linii różnych serii:

$$\tilde{\nu} = \frac{1}{\lambda} = R_H\left(\frac{1}{n^2} - \frac{1}{s^2}\right) \tag{30.33}$$

($\tilde{\nu}$ oznacza liczbę falową), przy czym stała Rydberga dla wodoru R_H wyraża się wzorem

$$R_H = \frac{2\pi^2 me^4}{h^3 c}.$$

Przejście do pierwiastków o większej liczbie atomowej Z wywołuje zmianę stałej Rydberga z R_H na R_Z. Uwzględnienie oddziaływania elektrostatycznego między elektronem e i ładunkiem jądra Ze prowadzi do pojawienia się dodatkowego czynnika Z^2, tak że

$$R_Z = Z^2 R_H.$$

A zatem biorąc pod uwagę wzór (30.33) stwierdzimy, że *ze wzrostem Z długości fal odpowiadających analogicznym przeskokom elektronowym* (np. z orbity drugiej na pierwszą — pierwsza linia serii Lymana w wodorze, lub z orbity trzeciej na drugą — pierwsza linia serii Balmera w wodorze) *maleją jak* Z^{-2}. Obliczając np. długość fali pierwszej linii „serii Balmera" dla platyny ($Z = 78$) znajdujemy rząd wielkości 10^{-1} nm, tzn. przesuwamy się z dziedziny promieniowania widzialnego do dziedziny promieniowania rentgenowskiego. Innymi słowy, w pierwiastkach ciężkich przeskoki elektronowe między wewnętrznymi powłokami elektronowymi, a więc np. z L do K, z M do K, z M do L itd. prowadzą do powstawania promieni Röntgena. Nasuwa się jednak pytanie dodatkowe, w jaki sposób mogą się odbywać przeskoki elektronowe tego rodzaju, skoro powłoki wewnętrzne atomów pierwiastków ciężkich są całkowicie zapełnione: powłoka K ma swą pełną obsadę w postaci dwóch elektronów, powłoka L — w postaci 8 itd. Przeskok elektronowy do powłoki K, L itd. może nastąpić tylko wtedy, gdy poprzedzony jest usunięciem z tych powłok elektronu pod wpływem działania jakiegoś czynnika zewnętrznego. Może się zdarzyć, że w lampie rentgenowskiej część rozpędzonych elektronów dobiegających do powierzchni antykatody wybije elektrony z wewnętrznych powłok atomowych, czyli spowoduje głęboką jonizację atomu polegającą na usunięciu nie elektronów wartościowości, lecz elektronów z głębszych powłok. Przeskoki z wyższych powłok na opróżnione miejsce, np. w powłoce K, będą prowadziły do powstawania linii rentgenowskich serii K. Linii K_α odpowiada przeskok z L na K, linii K_β — przeskok z M na K, linii K_γ — przeskok z N na K itd. Odpowiednio przeskoki na opróżnione miejsce w powłoce L prowadzą do powstania linii serii L itd. Zależność $\sqrt{\tilde{\nu}}$, a pośrednio długości fal linii różnych serii od liczby atomowej pierwiastka widoczna jest z wykresu (rys. 30.18).

W miarę wzrostu liczby atomowej Z analogiczne linie poszczególnych serii przesuwają się stopniowo w stronę fal krótszych. Dla danego Z linie serii K są najkrótsze. Linie powstające w opisany wyżej sposób tworzą widmo rentgenowskie, zwane charakterystycznym. Zmieniając antykatodę otrzymujemy widmo charakterystyczne o innych długościach fal.

Badaniem serii K różnych ciężkich pierwiastków zajmował się Moseley, który w r. 1913 ustalił następującą zależność między długością fali (liczbą falową) linii K_α i liczbą atomową pierwiastka:

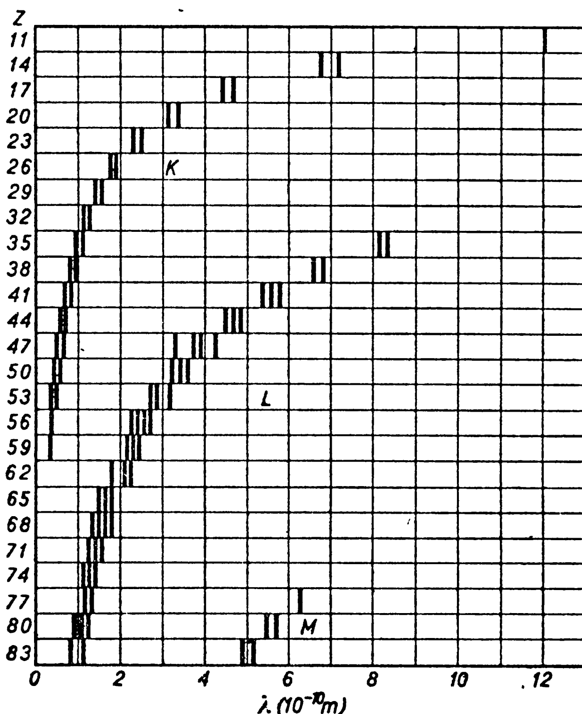
$$\sqrt{\frac{1}{\lambda R}} = 0{,}86(Z-1) \quad \text{lub} \quad \sqrt{\frac{\tilde{\nu}}{R}} = 0{,}86(Z-1), \tag{30.34}$$

gdzie R oznacza stałą Rydberga, a $\tilde{\nu} = \frac{1}{\lambda}$ — liczbę falową.

Przekształcając wzór Bohra dla serii widmowych (wzór (30.33)) przy założeniu, że $n = 1$, $s = 2$ (co odpowiada linii K_α) można go sprowadzić do postaci

$$\sqrt{\frac{1}{\lambda R}} = \sqrt{\frac{3}{4}} Z. \tag{30.35}$$

Porównajmy wzory (30.34) i (30.35). Lewe strony są równe, współczynniki liczbowe po prawej stronie można uważać za równe, niezgodność wzorów wiąże się z drugim czynnikiem po prawej stronie, a mianowicie z doświadczeń Moseleya wynika czynnik $(Z-1)$, a z teorii Bohra — czynnik Z. Jak wiadomo, liczba Z wiąże się z ładunkiem



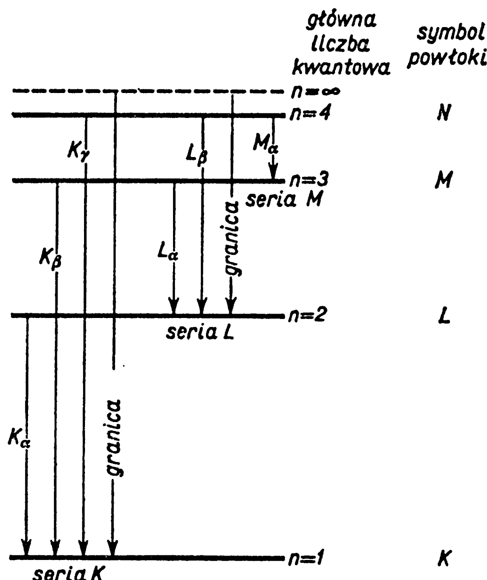
Rys. 30.18

dodatnim jądra i wyraża liczbę protonów w jądrze. Występowanie czynnika $(Z-1)$ we wzorze Moseleya można interpretować w ten sposób, że w czasie emisji kwantu odpowiadającego linii K_α jądro pierwiastka o liczbie atomowej Z oddziałuje na otoczenie tak, jakby liczba protonów w nim zawartych była o jeden mniejsza. Znajduje to swoje uzasadnienie w tym, że w atomie dowolnego pierwiastka, z wyjątkiem wodoru, w powłoce K są dwa elektrony. Jeżeli jeden z nich zostanie usunięty, to drugi, umieszczony stosunkowo bardzo blisko jądra, osłabi jego oddziaływanie na dalsze otoczenie. W przypadku linii K_α osłabienie to sprowadza się właśnie do wystąpienia we wzorze czynnika $(Z-1)$.

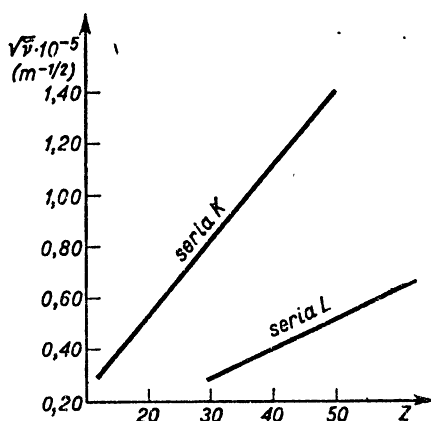
Dla dowolnej linii promieniowania rentgenowskiego prawo Moseleya można wyrazić jako

$$\sqrt{\frac{\tilde{\nu}}{R}} = a(Z-b), \tag{30.36}$$

gdzie a oznacza wartość stałą charakterystyczną dla danej linii danej serii, a b — tzw. *stałą przesłaniania*, wartość stałą charakterystyczną dla danej serii. Z danych doświadczalnych wynika, że dla linii serii K $b = 1$, dla linii serii L $b = 7{,}4$ itd.



Rys. 30.19



Rys. 30.20

Odpowiednikiem wzoru Bohra dla dowolnej linii dowolnej serii rentgenowskiej jest następujący wzór:

$$\tilde{\nu} = \frac{1}{\lambda} = (Z-b)^2 R\left(\frac{1}{n^2} - \frac{1}{s^2}\right).$$

Na rysunku 30.19 przedstawiono schematycznie mechanizm powstawania serii rentgenowskich za pośrednictwem poziomów energetycznych w atomie. Widać z niego wyraźnie, że liniom określonej serii odpowiada przeskok elektronu z dowolnego poziomu na poziom o określonej liczbie kwantowej głównej ($n = 1$ — seria K, $n = 2$ — seria L itd.). Wiemy już jednak z poprzednich rozważań, że poziomy energetyczne o liczbach $n > 1$ są rozszczepione energetycznie na podpoziomy (wartości $n = 2$ odpowiada rozszczepienie na 3 podpoziomy, $n = 3$ odpowiada 5 podpoziomom itd.). Stąd wniosek, że linie rentgenowskie mają również złożoną strukturę multipletową.

Prawo Moseleya przedstawione w postaci wzoru (30.36) wskazuje na liniową zależność między pierwiastkiem kwadratowym z liczby falowej promieni charakterystycznych rentgenowskich i liczbą atomową Z pierwiastka. Odkładając na osi rzędnych

$\sqrt{\tilde{\nu}}$, a na osi odciętych — Z otrzymuje się linie proste odpowiadające poszczególnym seriom (rys. 30.20). Prawidłowość przebiegu linii Moseleya byłaby naruszona, gdyby liczba atomowa pierwiastka w układzie okresowym wzrastała według rosnących mas atomowych, a nie wyrażała liczby protonów w jądrze. Odchylenia od prostoliniowego przebiegu obserwowalibyśmy przy następujących parach pierwiastków:

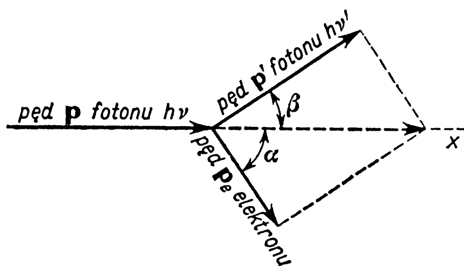
Liczba atomowa	Masa atomowa		Liczba atomowa	Masa atomowa
18	39,94	argon — potas	19	39,10
27	58,97	kobalt — nikiel	28	58,68
52	127,5	tellur — jod	53	126,92
90	232,05	tor — protaktyn	91	231

W wymienionych przypadkach wzrost masy atomowej nie idzie w parze ze wzrostem liczby protonów w jądrze, a zatem i ładunku jądra.

Prawo Moseleya miało podstawowe znaczenie dla fizyki atomowej. Liniowa zależność $\sqrt{\tilde{\nu}}$ od Z przyczyniła się do poprawnego określenia znaczenia liczby Z oraz do ustalenia liczby wolnych miejsc, które za czasów Moseleya istniały jeszcze w układzie okresowym pierwiastków.

30.13. Zjawisko Comptona

Dalsze badania właściwości promieni Röntgena doprowadziły w roku 1922 do wykrycia interesującego zjawiska, zwanego zjawiskiem Comptona. Podczas naświetlania grafitu twardymi promieniami Röntgena o częstotliwości ν Compton stwierdził istnienie w promieniowaniu rozproszonym fal o częstotliwości $\nu' < \nu$. Próby wyjaśnienia tego zjawiska wykazały, że jest to niemożliwe na gruncie teorii falowej. Podobnie jak



Rys. 30.21

w przypadku zjawiska fotoelektrycznego zachodzi konieczność traktowania promieniowania rentgenowskiego jako kwantów energii, którym można przypisywać pewną masę i pęd.

Zjawisko Comptona przedstawia schematycznie rys. 30.21. Kwant promieniowania $h\nu$ (foton rentgenowski) pada wzdłuż osi x na spoczywający elektron. W wyniku sprę-

żystego zderzenia elektron odrzucony zostaje z prędkością v pod kątem α względem osi x, a równocześnie tor fotonu odchyla się o kąt β w przeciwną stronę. Foton oddaje przy zderzeniu część swej energii, a więc po zderzeniu ma mniejszą częstotliwość ν', a dłuższą falę. Przy zderzeniu sprężystym obowiązuje zasada zachowania energii i pędu, czyli

$$h\nu = h\nu' + \tfrac{1}{2}mv^2, \qquad \mathbf{p} = \mathbf{p}' + \mathbf{p}_e,$$

gdzie $\mathbf{p}$ oznacza wektor pędu fotonu przed zderzeniem, $\mathbf{p}'$ — wektor pędu fotonu po zderzeniu, $\mathbf{p}_e$ — wektor pędu elektronu. Uwzględnienie wzorów teorii względności prowadzi do następującego wyrażenia na zmianę długości fali $\Delta\lambda$ promienia rozproszonego:

$$\Delta\lambda = \frac{h}{m_0 c}(1-\cos\alpha).$$

Zgodność wniosków teoretycznych z wynikami doświadczalnymi potwierdza słuszność założeń, w których promieniom rentgenowskim przypisywaliśmy cechy korpuskularne: określoną masę i określony pęd. Trzeba jednak podkreślić, że zjawisko Comptona zachodzi wyraźnie tylko wtedy, gdy padające promienie mają dużą częstotliwość i tylko w przypadku zderzeń fotonów z elektronami słabo związanymi z atomami, gdyż jedynie wtedy można je uważać za swobodne. Gdy elektron jest silnie związany z atomem o dużej masie, nie zachodzi wymiana pędów i energii: rozproszone fotony mają niezmienioną częstotliwość. Na ogół w rozproszonym promieniowaniu rentgenowskim występują promienie zarówno o zmienionej, jak i niezmienionej częstotliwości, z przewagą tych pierwszych w przypadku lekkich, rozpraszających atomów.

Pytania i zadania

1. Jakie wnioski wynikają z doświadczeń nad przenikaniem cząstek α przez materię?
2. Czym się różnią modele budowy atomu Rutherforda i Rutherforda–Bohra?
3. Co charakteryzuje tory stacjonarne elektronów w atomie? Czy ujęcie Bohra jest zgodne z elektrodynamiką klasyczną?
4. Jak się przedstawia widmo wodoru?
5. Wyjaśnić mechanizm emisji promieniowania przez wodór na podstawie modelu Bohra.
6. Wyprowadzić na podstawie modelu Bohra następujące wzory na: 1) promień orbity elektronu, 2) prędkość na orbicie, 3) energię całkowitą elektronu na orbicie, 4) liczbę falową związaną z przeskokiem elektronowym z orbity s na orbitę n.
7. Jakie fakty doświadczalne nie mogły być wyjaśnione na podstawie modelu Bohra? Jakie poprawki i uzupełnienia wprowadzono do modelu Bohra?
8. Ile liczb kwantowych charakteryzuje elektron w atomie? Jaki jest ich sens fizyczny?
9. Co mówi zasada Pauliego?
10. Jak można tłumaczyć rozbudowę powłok elektronowych kolejnych pierwiastków układu Mendelejewa?
11. Jaka jest maksymalna liczba elektronów odpowiadająca liczbie kwantowej głównej $n = 4$?
12. Co oznaczają symbole K, L, M itp. i s, p, d itp.?
13. Co to jest moment magnetyczny orbitalny elektronu? Co to jest magneton Bohra?
14. Jaką hipotezę wysunął de Broglie? Jakie fakty doświadczalne potwierdziły tę hipotezę?

15. W jakim przyrządzie wykorzystuje się naturę falową elektronu? Podać zasadę budowy, działania i zastosowania tego przyrządu.

16. Jak można tłumaczyć powstawanie: a) widma ciągłego promieniowania rentgenowskiego, b) widma charakterystycznego promieniowania rentgenowskiego, c) granicy krótkofalowej widma ciągłego promieniowania rentgenowskiego.

17. Wyjaśnić mechanizm powstawania serii widmowych promieniowania rentgenowskiego.

18. Jaka jest treść prawa Moseleya? Jakie konsekwencje wypływają z tego prawa?

19. Obliczyć promienie drugiej i trzeciej orbity w atomie wodoru oraz prędkości elektronów na tych orbitach.

20. Obliczyć całkowitą energię elektronu na orbicie drugiej i trzeciej w atomie wodoru. Wyrazić te energie w J i eV.

21. Obliczyć długość fali czwartej linii serii Balmera.

Odp. $4,1 \cdot 10^{-7}$ m.

22. Obliczyć energię jonizacji atomu wodoru.

Odp. $21,85 \cdot 10^{-19}$ J.

23. Obliczyć moment orbitalny magnetyczny elektronu w atomie wodoru na orbicie pierwszej.

Odp. $9,1 \cdot 10^{-24}$ $A \cdot m^2$.

24. Jaki jest stosunek momentu magnetycznego elektronu na dowolnej orbicie atomu wodoru do momentu pędu elektronu na tej samej orbicie?

Odp. $e/2m$.

25. Jaką długość fali de Broglie'a mają elektrony, których energia kinetyczna wynosi 100 eV?

Odp. $1,2 \cdot 10^{-10}$ m.

26. Foton promieniowania rentgenowskiego o długości fali $\lambda = 1$ Å trafia na słabo związany elektron atomu pierwiastka lekkiego i odchyla się od pierwotnego biegu o kąt 90°. Obliczyć długość fali fotonu rozproszonego.

Odp. 1,024 Å.

27. Jaką energię kinetyczną uzyska uderzony elektron przy spełnieniu warunków poprzedniego zadania?

Odp. $4,7 \cdot 10^{-17}$ J.

ROZDZIAŁ 31

Fizyka jądrowa

31.1 Ogólna charakterystyka jądra atomowego

Jak już wspominaliśmy w § 30.2, doświadczenia Geigera i Marsdena, dotyczące rozpraszania cząstek α przy ich przechodzeniu przez cienkie folie metalowe, doprowadziły do utworzenia modelu atomu Rutherforda. Według tego modelu prawie całkowita masa atomu i jego ładunek dodatni skupione są w wewnętrznej części atomu, w tzw. jądrze atomowym, o promieniu 10^{-14}–10^{-15} m, mniejszym około 10 tysięcy razy od promienia atomu jako całości. Późniejsze koncepcje modelowej struktury atomu (Bohr, Sommerfeld) utrzymały założenie istnienia jądra.

Badania fizyki współczesnej przyniosły dużo wiadomości o jądrze atomowym. Po odrzuceniu błędnej hipotezy, że jądra atomowe składają się z protonów i elektronów, przyjmuje się obecnie, zgodnie z sugestią Heisenberga i Iwanienki, że jądra atomowe złożone są z tzw. nukleonów, tzn. protonów i neutronów. *Protony* są to cząstki elementarne o dodatnim ładunku elementarnym $e = 1{,}6021 \cdot 10^{-19}$ C i masie spoczynkowej $m_p = 1{,}6725 \cdot 10^{-27}$ kg. Jest to masa około 1836 razy większa od masy spoczynkowej elektronu $m_e = 9{,}109 \cdot 10^{-31}$ kg. Pojedynczy proton stanowi jądro wodoru. *Neutrony* są to cząstki elementarne, elektrycznie obojętne, o masie spoczynkowej $m_n = 1{,}6748 \times \times 10^{-27}$ kg, około 1839 razy większej od masy spoczynkowej elektronu.

Liczba protonów i neutronów w jądrze charakteryzuje atom pierwiastka, a mianowicie liczba protonów Z, zwana *liczbą atomową*, określa ładunek dodatni jądra (jako wielokrotność ładunku elementarnego e) i jest równocześnie numerem porządkowym odpowiedniego pierwiastka w układzie okresowym Mendelejewa, a suma liczby protonów i neutronów (czyli ogólna liczba nukleonów) w jądrze przedstawia tzw. *liczbę masową A*. Każdy pierwiastek może być opisany przez podanie jego symbolu X oraz liczby atomowej Z i masowej A. Stosowane są różne postacie zapisu, np. ${}^{A}_{Z}X$, ${}_{Z}X^{A}$. Można też posługiwać się zapisem ${}^{A}X$, np. ${}^{12}C$, ${}^{238}U$, gdyż symbol chemiczny pierwiastka określa jednoznacznie jego liczbę porządkową w układzie okresowym, a tym samym liczbę atomową Z.

Analiza produktów rozpadu promieniotwórczego oraz dokładne badania mas atomowych różnych pierwiastków (Thomson, Aston i inni) doprowadziły do wykrycia

izotopii. Polega ona na tym, że atomy danego pierwiastka mogą występować w różnych odmianach, przy czym niezmienna jest liczba protonów, czyli zachowany jest ładunek dodatni jądra, a zmianie ulega liczba neutronów (a w związku z tym liczba masowa A). Jako przykłady izotopii rozpatrzymy izotopy wodoru i uranu. Wodór występuje w trzech odmianach izotopowych, a mianowicie jako: 1) ${}^{1}_{1}H$ — wodór normalny (nazwa rzadko stosowana — prot) stanowiący 99,98% całego wodoru występującego w przyrodzie, 2) ${}^{2}_{1}H$ — wodór ciężki, tzw. *deuter*, stanowiący blisko 0,02% oraz 3) ${}^{3}_{1}H$ — tzw. tryt — odmiana spotykana bardzo rzadko, jak wynika z podanych poprzednio procentów. Jądrom tych izotopowych odmian wodoru nadano nazwy *proton*, *deuteron* i *tryton*. Odmiany izotopów uranu w podobnym zestawieniu przedstawiają się następująco:

${}^{234}_{92}U$ — bardzo rzadki,

${}^{235}_{92}U$ — około 0,7%,

${}^{238}_{92}U$ — około 99,3%,

${}^{239}_{92}U$ — otrzymywany na drodze sztucznych przemian jądrowych.

O odmianach izotopowych uranu będziemy mówili szerzej w § 31.7 i 31.8.

Warto podkreślić, że zjawisko izotopii jest bardzo rozpowszechnione. Nieliczne pierwiastki (około 20) występują tylko w jednej postaci. Ogólna liczba poznanych dotychczas izotopów (zarówno trwałych, jak i promieniotwórczych) przewyższa 1500.

31.2. Trwałość jądra

Na zagadnienie trwałości jąder atomowych spojrzymy teraz z punktu widzenia: a) energii wiązania, b) sił działających w jądrze, c) możliwych przemian prowadzących do wytworzenia jąder o większej trwałości.

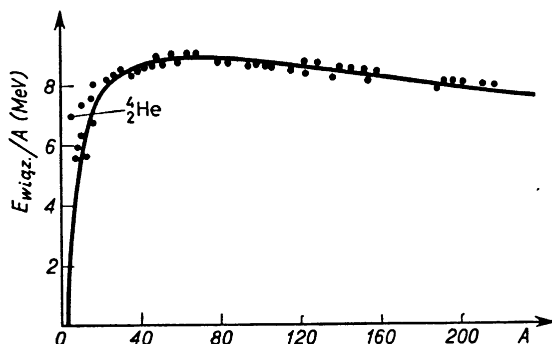
31.2.1. *Energia wiązania*. Wydawałoby się, że dokładne ustalenie masy jądra nie nastręcza żadnych trudności, skoro znamy liczbę protonów i neutronów w jądrze oraz ich masy m_p i m_n. Okazuje się jednak, że masa jądra nie równa się sumie mas składników, jest zwykle mniejsza od sumy mas składników. W związku z tym mówimy o tzw. *defekcie* (niedoborze) *masy jądra*. Defekt masy $\Delta m = (\sum m_{\text{składników}} - m_{\text{jądra}})$ może być łatwo przeliczony na energię, zgodnie z wzorem Einsteina (7.9):

$$E = \Delta mc^2.$$

Wynik uzyskany z takiego obliczenia przedstawia energię, która wyzwala się podczas łączenia się odpowiedniej liczby nukleonów w jądro atomowe. Energię tę nazywamy *energią wiązania jądra*. Wyrażamy ją zwykle w megaelektronowoltach (MeV)*. Dzieląc energię wiązania jądra przez liczbę tworzących go nukleonów (równą liczbie masowej A) znajdujemy energię wiązania jednego nukleonu, tzn. ilość energii wyzwalanej podczas przyłączania jednego nukleonu do jądra lub energii, jakiej należy dostarczyć, aby jeden

* 1 MeV = 1,6019 · 10^{-13} J.

nukleon usunąć z jądra. Rysunek 31.1 przedstawia zależność energii wiązania nukleonu od liczby masowej A. W odniesieniu do pierwiastków o małej wartości A warto zwrócić uwagę na dużą wartość energii wiązania nukleonu $E = 7$ MeV w jądrze helu. Następnie charakterystyczne jest utrzymywanie się dużej wartości energii wiązania nukleonu dla pierwiastków o liczbach masowych $40 < A < 140$. Można przyjmować, że średnia wartość tej maksymalnej energii wiązania nukleonu wynosi 8,5 MeV. Obniżenie wartości energii wiązania nukleonu jest charakterystyczne dla najcięższych pierwiastków układu okresowego, dla których wartość ta spada do około 7,5 MeV. W związku z tym nasuwa się interesująca możliwość: gdyby jądro ciężkie, np. uranu lub toru, rozszczepiło się (pękło) na dwa jądra o liczbach A zawartych w granicach 40–140, to wobec większej energii wiązania tych nowych jąder atomowych (przeciętnie o 1 MeV na nukleon) wyzwoliłaby się przy tym olbrzymia ilość energii około 200 MeV. Zobaczymy, że nie jest to tylko hipoteza; „pękania" jąder ciężkich są znanym od r. 1939 faktem, o którym szerzej będzie mowa w § 31.7.



Rys. 31.1

31.2.2. *Siły działające w jądrach atomowych.* Na zagadnienie trwałości jądra atomowego można spojrzeć od strony działających w nim sił. Rozróżniamy tu dwa rodzaje sił, a mianowicie siły: a) jądrowe oraz b) elektrostatyczne.

Siły jądrowe — o nie ustalonej jeszcze dokładnie naturze — są siłami przyciągania małego zasięgu, działającymi między wszelkimi składnikami jądra, a więc istnieją oddziaływania jądrowe $p \to p$, $n \to n$, $p \to n$. Siły te osiągają znaczne wartości przy wzajemnych odległościach cząstek od $0{,}5 \cdot 10^{-15}$ m do 10^{-15} m. W przypadku odległości większych od 10^{-15} m siły jądrowe gwałtownie spadają do zera.

Siły elektrostatyczne (kulombowskie) działające między protonami wchodzącymi w skład jądra są siłami dalekiego zasięgu, odwrotnie proporcjonalnymi do kwadratu odległości.

Uwzględnienie sił działających w jądrze pozwala na ustalenie optymalnych warunków trwałości jądra. Rozważania teoretyczne prowadzą do wniosku, że gdyby w jądrze nie występowały siły odpychania elektrostatycznego, to najtrwalsze byłyby jądra o równej liczbie protonów i neutronów, tzn. takie, dla których

$$Z = A - Z,$$

czyli
$$Z = \frac{A}{2}.$$

W warunkach rzeczywistych istniejących w jądrach protony są elementem „destrukcyjnym". Aby więc jądro było trwałe, powinno ono zawierać mniej protonów niż neutronów, czyli Z powinno być mniejsze od $A/2$. Zaznacza się to tym wyraźniej, im jądro jest cięższe. Wyjątkową pozycję wśród jąder trwałych mają ${}^{1}_{1}H$ i ${}^{3}_{2}He$, w których istnieje nadmiar protonów w stosunku do neutronów. W przypadku nielicznych, również raczej lekkich, pierwiastków trwałość budowy idzie w parze ze spełnieniem warunku $Z = A/2$. Można tu przykładowo wymienić: ${}^{4}_{2}He$, ${}^{6}_{3}Li$, ${}^{12}_{6}C$, ${}^{16}_{8}O$, ${}^{20}_{10}Ne$.

W przyrodzie występują najczęściej jądra o parzystej liczbie protonów i neutronów, tzw. jądra parzysto-parzyste (162 odmiany), najrzadziej — jądra nieparzysto-nieparzyste (4 odmiany).

Jądra ciężkie o dużej zawartości protonów począwszy od $Z > 81$ i $A > 210$ są nietrwałe, ulegają naturalnemu rozpadowi promieniotwórczemu, a pierwiastki o liczbie $Z > 90$ podlegają nawet pękaniu. Brak też w przyrodzie trwałych *transuranów*, tzn. pierwiastków o liczbie $Z > 92$, które można otrzymać tylko na drodze sztucznych przemian.

31.2.3. *Wstępne zestawienie przemian jądrowych.* W zależności od tego, czy dane jądro w porównaniu z optymalnymi warunkami trwałości ma nadmiar protonów czy neutronów, wykazuje ono tendencję do przemiany jednej postaci nukleonu w drugą. Równanie tych przemian napiszemy na razie w postaci uproszczonej (por. punkt 31.3.4):


$$p^+ \to n + e^+ \qquad n \to p^+ + e^-,$$

tzn. proton przechodzi w neutron z równoczesną emisją pozytonu e^+ lub neutron przechodzi w proton z równoczesną emisją negatonu e^-. Emisja negatonu lub pozytonu musi towarzyszyć tym przemianom, gdyż inaczej nie byłaby spełniona zasada zachowania ładunku elektrycznego. Negaton jest to dobrze nam już znany elektron o ładunku elementarnym ujemnym, a pozyton jest to antyelektron, cząstka o masie równej masie elektronu i ładunku elementarnym dodatnim (pozyton wykryty został w r. 1932 przez Andersona podczas badań promieni kosmicznych). W przypadku wyżej wymienionych przemian w jądrze atomowym mówimy o promieniotwórczości β^+ (emisja pozytonu — § 31.5) i β^- (emisja elektronu — punkt 31.3.1).

Przemiana protonu w jądrze na neutron może się także dokonać przez wchłonięcie do jądra elektronu z jednej z powłok zewnętrznych najbliższych jądra, np. z powłoki K lub L. Reakcja tego typu zwana jest *przemianą K* (lub *wychwytem K*, *K — capture*) lub odpowiednio przemianą L.

Przemiana jądrowa prowadząca do korzystniejszego z punktu widzenia trwałości jądra stosunku liczby protonów i neutronów może także nastąpić w wyniku reakcji rozszczepienia, podczas której ciężkie jądro pęka na dwa lżejsze fragmenty, najczęściej z równoczesną emisją neutronów. O tym typie przemiany jądrowej mówimy w § 31.7.

Dla uzupełnienia tego zestawienia przemian jądrowych warto jeszcze wspomnieć o promieniotwórczości α (por. p. 31.3) oraz o promieniotwórczości protonowej sztucznie otrzymanych jąder atomowych.

31.3. Promieniotwórczość naturalna

Omawianie przemian jądrowych rozpoczniemy od zjawiska promieniotwórczości naturalnej, ściśle związanego z pracami badawczymi Marii Skłodowskiej - Curie, Piotra Curie oraz Henryka Becquerela. Uczeni ci otrzymali za odkrycie promieniotwórczości nagrodę Nobla w r. 1903.

31.3.1. *Ogólna charakterystyka przemian promieniotwórczych naturalnych. Reguły przesunięć.* Nie uwzględniając chronologicznego porządku odkryć związanych z promieniotwórczością naturalną zestawimy najważniejsze wnioski:

1) Promieniotwórczość naturalna sprowadza się do przemian zachodzących w jądrach pierwiastków ciężkich, począwszy od $Z = 81$ do $Z = 92$.

2) Zmiany zachodzące w jądrach pierwiastków promieniotwórczych naturalnych dzielimy na przemiany α i β. Obu tym przemianom może towarzyszyć emisja promieni γ. *Przemiana* α sprowadza się do wyrzucenia z badanego jądra dodatnio naładowanej cząstki α $(2p+2n)$, natomiast przemiana β wiąże się z emisją elektronu ujemnego (negatonu $-e$). A zatem promieniowanie α i β ma charakter korpuskularny. Emisja cząstek α lub β może pozostawić *jądro w stanie wzbudzonym*, tzn. w stanie energetycznie wyższym od stanu normalnego. Przejściu jądra ze stanu wzbudzonego do stanu normalnego towarzyszy emisja krótkofalowego promieniowania γ. Długości fal elektromagnetycznego promieniowania γ sięgają od $5 \cdot 10^{-13}$ m do $4 \cdot 10^{-11}$ m (por. przegląd fal elektromagnetycznych § 23.7).

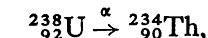
Różny charakter elektryczny cząstek α i β oraz odmienność charakteru promieni γ wykazują znane doświadczenia, polegające na zastosowaniu pola elektrycznego lub magnetycznego. Promienie emitowane z preparatu promieniotwórczego R przechodząc przez pole elektryczne między okładkami kondensatora rozdzielają się na trzy wiązki: jedna odchyla się w stronę okładki ujemnej, druga — w stronę okładki dodatniej, trzecia — biegnie prostoliniowo bez zakłóceń ze strony pola. Pierwszą wiązkę stanowią cząstki α, drugą — cząstki β, czyli elektrony, natomiast trzecią — promienie γ o naturze fali elektromagnetycznej. Analogiczne wyniki można osiągnąć pod wpływem działania pola magnetycznego prostopadłego do kierunku pierwotnego biegu promieni (rys. 31.2). Promienie γ przechodzą znów bez odchylenia. Cząstki α i β odchylają się w przeciwnych kierunkach w płaszczyźnie prostopadłej do wektora $\mathbf{H}$ i wektora $\mathbf{v}$. Zwrot odchylenia wynika z wzoru na siłę (wzór 21.2).

Wymienione trzy rodzaje promieniowania różnią się bardzo pod względem przenikliwości przez materię. Na przykład cząstki α mogą z trudem przeniknąć przez 1 kartkę zwykłego papieru, cząstki β przechodzą przez 100 takich kartek, a do zatrzymania promieni γ trzeba wielu grubych tomów. Osłonę przed promieniami β stanowią blachy o grubości kilku milimetrów, a osłonę przed promieniami γ dają dopiero grube bloki ołowiane.

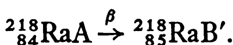
3) Emisja cząstki α zmniejsza ładunek dodatni jądra o dwie jednostki (liczba atomowa Z maleje o dwa) oraz jego masę o cztery jednostki (liczba masowa A maleje o cztery). A zatem wskutek przemiany α jądro danego pierwiastka przekształca się w jądro pierwiastka przesuniętego o dwa miejsca wstecz w układzie Mendelejewa.

4) Emisja cząstki β powiększa dodatni ładunek jądra o jednostkę (liczba atomowa Z rośnie o jeden) nie zmieniając liczby masowej. A zatem wskutek przemiany β jądro danego pierwiastka przekształca się w jądro pierwiastka przesuniętego o jedno miejsce naprzód w układzie Mendelejewa.

5) Prawidłowości ujęte w punktach 3 i 4, znane pod nazwą *prawa przesunięć Fajansa i Soddy'ego,* stwierdzone były w latach 1911–1913, gdy jeszcze nieznana była budowa jądra atomu. Ilustracją prawa przesunięć mogą być następujące reakcje:

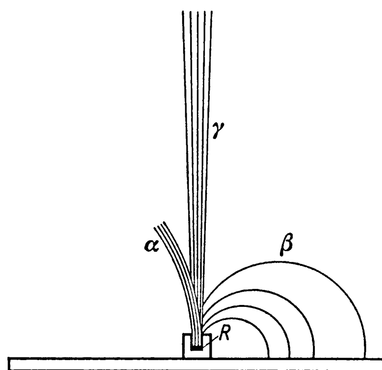


$$^{238}_{92}U \xrightarrow{\alpha} {}^{234}_{90}Th,$$



$$^{218}_{84}RaA \xrightarrow{\beta} {}^{218}_{85}RaB'.$$

Symbole nad strzałką oznaczają rodzaj zachodzącej przemiany.



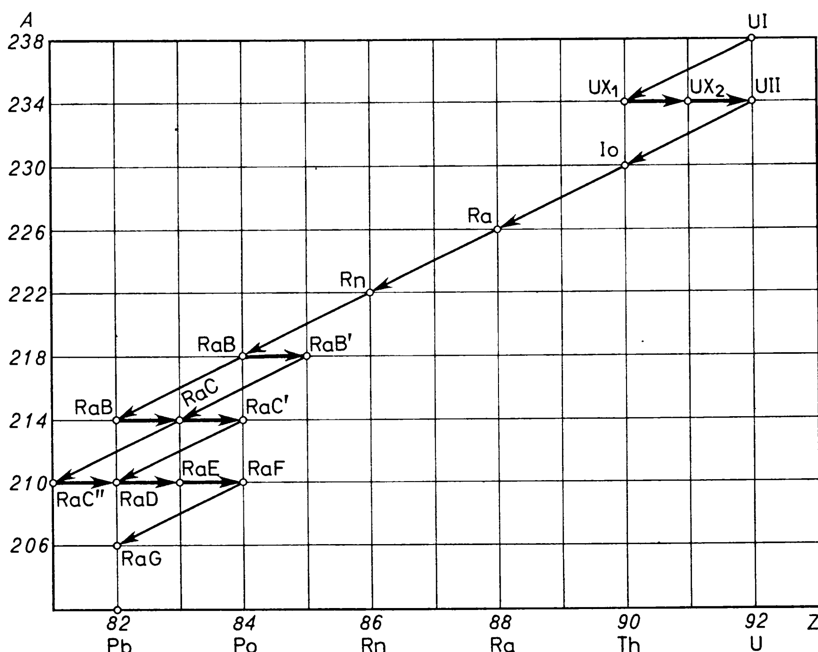
Rys. 31.2

31.3.2. *Szeregi (rodziny) promieniotwórcze.* Pierwiastki ciężkie, wykazujące promieniotwórczość naturalną, można zgrupować w następujące szeregi promieniotwórcze: urano-radowy, urano-aktynowy, torowy, neptunowy. Ostatni spośród wymienionych szeregów obejmuje oprócz pierwiastków występujących w przyrodzie pewną liczbę otrzymanych sztucznie transuranów, takich jak neptun ($Z = 93$), pluton ($Z = 94$) i ameryk ($Z = 95$).

Na rys. 31.3 przedstawione są schematycznie przemiany promieniotwórcze w szeregu urano-radowym. Na osi odciętych odłożone są liczby atomowe Z, na osi rzędnych — liczby masowe A. Szereg rozpoczyna się od $^{238}_{92}U$. Strzałka ukośna kończąca się na $^{234}_{90}UX_1$ przedstawia przemianę α, następne dwie kolejne strzałki poziome, kończące się odpowiednio na $^{234}_{91}UX_2$ i $^{234}_{92}UII$, odpowiadają zgodnie z prawem Fajansa-Soddy'ego dwóm przemianom β. Dalej odbywają się cztery kolejne przemiany α, z tym że jednym z pośrednich ogniw szeregu jest $^{226}_{88}Ra$. Interesujące są pierwiastki promieniotwórcze $^{218}_{84}RaA$ oraz $^{214}_{83}RaC$. Jądra obu wymienionych pierwiastków mogą podlegać zarówno przemianie α, jak i β, a więc jak widać z rysunku, produktami rozpadu RaA mogą być zarówno $^{214}_{82}RaB$, jak i $^{218}_{85}RaB'$. Rozwidlenie szeregu zostaje jednak szybko zlikwidowane, gdyż RaB podlega przemianie β, a RaB' przemianie α,

wskutek czego otrzymujemy w obu przypadkach $^{214}_{83}$RaC. Analogiczne rozwidlenie mające początek w RaC zostaje zlikwidowane w RaD, który po dwóch kolejnych przemianach β i jednej przemianie α przechodzi w końcowy, trwały pierwiastek układu Mendelejewa, a mianowicie RaG, będący izotopem ołowiu $^{206}_{82}$Pb.

Spoglądając na wykres można łatwo dostrzec, że przemiany promieniotwórcze stosunkowo często prowadzą do powstawania izotopów. Na przykład izotopami uranu są: $^{238}_{92}$UI i $^{234}_{92}$UII, izotopami polonu są: $^{218}_{84}$RaA, $^{214}_{84}$RaC′ i $^{210}_{84}$RaF, a izotopami ołowiu są: $^{214}_{82}$RaB, $^{210}_{82}$RaD i $^{206}_{82}$RaG.



Rys. 31.3

Łatwo też sprawdzić, że w każdym z szeregów promieniotwórczych liczba masowa A w czasie kolejnej przemiany albo w ogóle nie ulega zmianie (przemiana β), albo zmienia się o 4 jednostki (przemiana α). Jeśli zatem w rozważanym szeregu urano-radowym liczbę A dla początkowego ogniwa $^{238}_{92}$UI przedstawimy jako

$$A = 238 = 4\cdot 59+2,$$

to jedynie możliwymi dla pozostałych ogniw szeregu będą wartości $4\cdot 58+2$, $4\cdot 57+2$ itd. do $4\cdot 51+2 = 206$ (ogniwo końcowe $^{206}_{82}$Pb, czyli RaG). Innymi słowy, liczbę masową tego szeregu można przedstawić jako

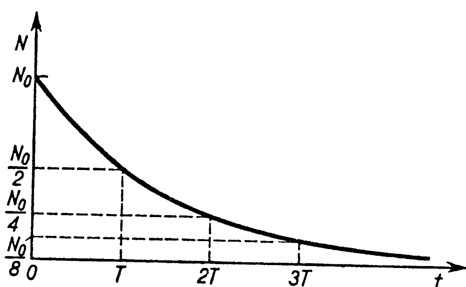
$$A = 4n+2, \quad \text{gdzie} \quad 51 \leqslant n \leqslant 59.$$

Rozważając w analogiczny sposób liczby masowe odpowiadające pozostałym szeregom promieniotwórczym stwierdzono, że w szeregu urano-aktynowym $A = 4n+3$ (ogniwo

początkowe aktyno-uran $^{235}_{92}$AcU, końcowe $^{207}_{82}$Pb; $51 \leqslant n \leqslant 58$), w szeregu torowym $A = 4n$ (ogniwo początkowe tor $^{232}_{90}$Th, końcowe $^{208}_{82}$Pb; $52 \leqslant n \leqslant 58$), w szeregu neptunowym $A = 4n+1$ (ogniwo początkowe pluton $^{241}_{84}$Pu, końcowe tal $^{205}_{81}$Tl; $51 \leqslant n \leqslant 60$). Warto podkreślić, że w szeregu neptunowym ostatnim ogniwem nie jest już izotop ołowiu, jak również, że w tym szeregu nie występuje produkt gazowy. W pozostałych szeregach takie produkty występują. Są to tzw. emanacje, którym nadano nazwy radonu, toronu i aktynonu.

Promieniotwórczość naturalna występuje też sporadycznie wśród pierwiastków lżejszych. Na przykład neodym ^{144}Nd, samar ^{147}Sm i platyna ^{190}Pt wykazują promieniotwórczość α, a potas ^{40}K, rubid ^{87}Rb, ind ^{115}In i jeszcze parę innych ulegają przemianie β. W tych przypadkach nie tworzą się jednak szeregi promieniotwórcze.

31.3.3. *Kinetyka przemian promieniotwórczych.* Rozważając szeregi promieniotwórcze warto chwilę uwagi poświęcić kinetyce przemian. Niech N oznacza ogólną liczbę



Rys. 31.4

jąder pierwiastka promieniotwórczego, a dN — liczbę jąder ulegających rozpadowi w czasie dt proporcjonalną do N i do dt, a zatem

$$-dN = \lambda N dt,$$

$$-\frac{dN}{N} = \lambda\, dt,$$

gdzie $\lambda = \text{const}$ oznacza *stałą rozpadu* o wartości charakterystycznej dla pierwiastka promieniotwórczego. Po scałkowaniu otrzymujemy zależność

$$\ln N = -\lambda t + \text{const}.$$

Zakładając warunki początkowe: w chwili $t = 0$, $N = N_0$, znajdujemy

$$\ln N_0 = \text{const},$$

a zatem

$$\ln\frac{N}{N_0} = -\lambda t,$$

czyli

$$N = N_0 e^{-\lambda t}.$$

Innymi słowy, liczba atomów pierwiastka promieniotwórczego maleje wykładniczo z biegiem czasu. Przedstawia to wykres na rys. 31.4.

Rozpad promieniotwórczy danego pierwiastka można też charakteryzować za pomocą tzw. okresu połowicznego zaniku T. *Okres połowicznego zaniku* jest to czas, po upływie którego początkowa liczba atomów zmniejsza się do połowy. A zatem dla $t = T$, $N = \frac{N_0}{2}$ i równanie (31.1) przyjmuje postać

$$\frac{N_0}{2} = N_0 e^{-\lambda T},$$

skąd

$$2 = e^{\lambda T},$$

$$T = \frac{\ln 2}{\lambda} \approx \frac{0,693}{\lambda}. \qquad (31.2)$$

Wyrażenie (31.2) wiąże okres połowicznego zaniku ze stałą rozpadu λ.

Tabela 31.1

Okresy połowicznego zaniku i stałe rozpadu

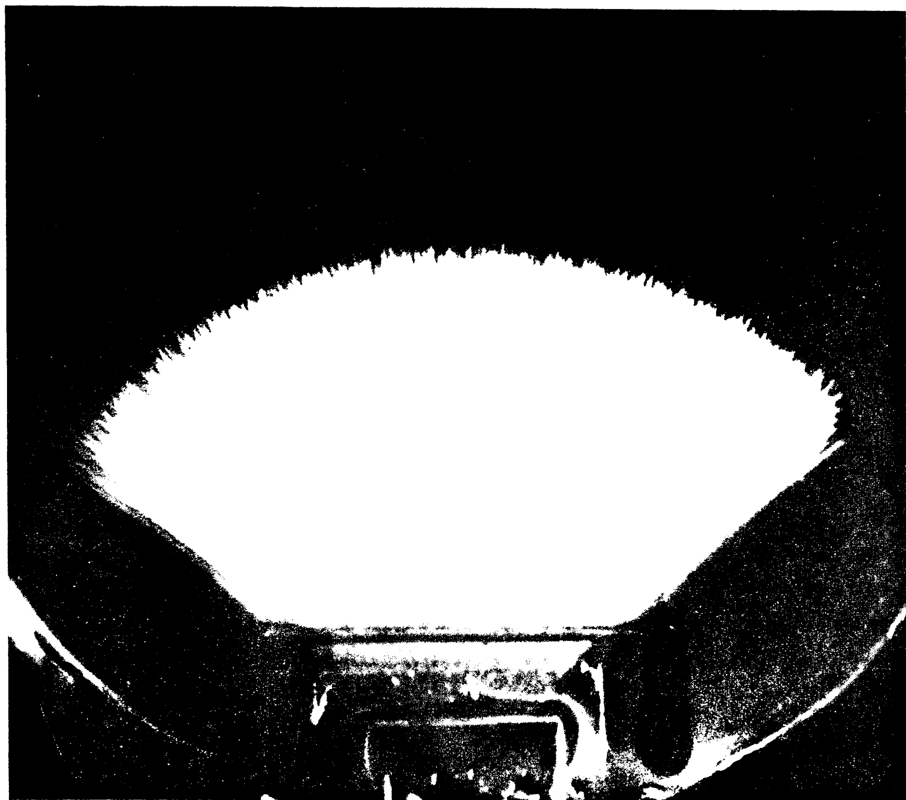
Substancja promieniotwórcza	Symbol	T	λ s^{-1}
uran I	^{238}U	$4,51 \cdot 10^9$ lat	$4,87 \cdot 10^{-18}$
rad	^{226}Ra	1620 lat	$1,36 \cdot 10^{-11}$
rad C′ (izotop polonu)	^{214}Po	$1,64 \cdot 10^{-4}$ s	$4,23 \cdot 10^3$
aktyn	^{227}Ac	21,8 lat	$1,01 \cdot 10^{-9}$
tor	^{232}Th	$1,39 \cdot 10^{10}$ lat	$1,58 \cdot 10^{-18}$

Wartości liczbowe okresów połowicznego zaniku (tab. 31.1) zmieniają się dla różnych pierwiastków w bardzo szerokich granicach, od mikrosekund do przeszło 10^{16} lat.

31.3.4. *Energia cząstek α i β. Neutrino.* Rozpad promieniotwórczy odbywa się zazwyczaj w powietrzu. Wiemy już, że podstawowym zjawiskiem towarzyszącym przechodzeniu promieni α i β przez powietrze jest jonizacja (por. punkt 20.9.1). Na zjonizowanie każdej napotkanej cząsteczki gazu cząstka α lub β traci część swej energii kinetycznej. Długość odcinka drogi, na którym cząstka danego rodzaju wywołuje jonizację w powietrzu w warunkach normalnych, czyli tzw. *zasięg cząstki* może więc być miernikiem początkowej energii kinetycznej, z jaką cząstka opuszcza jądro. Zbadanie długości zasięgów możliwe jest np. przy zastosowaniu komory Wilsona, której zasada działania opisana jest w § 30.2. Z fotografii torów cząstek α wylatujących z określonego preparatu promieniotwórczego (rys. 31.5) widać, że mają one jednakowe zasięgi. Innymi słowy, cząstki α wylatują z danego preparatu promieniotwórczego z jednakowymi energiami E_k, a więc z jednakowymi prędkościami. Prędkości cząstek α są rzędu 10^4 km/s.

Rozpatrzmy bardziej szczegółowo zagadnienie energii cząstek α emitowanych w czasie rozpadu promieniotwórczego. Niech z macierzystego jądra $^{A}_{Z}X$ w wyniku emisji α powstanie jądro $^{A-4}_{Z-2}Y$. Reakcji rozpadu towarzyszy wyzwolenie energii, którą otrzymuje wylatująca cząstka α oraz nowo powstałe jądro. Stosując dwie podsta-

wowe zasady zachowania; zasadę zachowania energii i pędu — można wykazać, że w przypadku ciężkich jąder o dużym A prawie cała energia rozpadu przypada w udziale emitowanej cząstce α. Tak więc wyznaczając energię cząstki α (np. z pomiaru zasięgu) znajdujemy prawie całkowitą energię rozpadu α.



Rys. 31.5

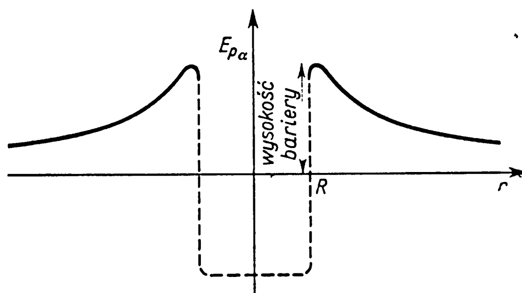
Reakcja rozpadu przebiega samorzutnie, bez udziału sił zewnętrznych, a więc energia powstaje kosztem zmniejszenia masy

$$E = (m_X - m_Y - m_\alpha)c^2.$$

Wyrażenie w nawiasie ma wartość dodatnią dla każdej samorzutnej emisji α, jest to bowiem warunek wystąpienia samorzutnego rozpadu. Może warto tu podkreślić, że w przyrodzie nie obserwuje się samorzutnej emisji neutronowej, protonowej i deuteronowej z jąder naturalnych występujących w przyrodzie, gdyż wartości zmian masy, a co za tym idzie i energii, są w tych przypadkach ujemne.

Wyznaczone doświadczalnie energie kinetyczne cząstek α (np. na podstawie pomiarów zasięgów) leżą w granicach 4,1 MeV dla $^{232}_{90}$Th i 8,8 MeV dla $^{212}_{84}$Po. I tu właśnie pojawia się problem nie do rozwiązania z punktu widzenia fizyki klasycznej. Z roz-

ważań dotyczących zależności potencjału od odległości od środka jądra $V = V(r)$, jak również zależności energii potencjalnej cząstki α od odległości wynika istnienie bariery potencjału i bariery energetycznej, otaczającej jądra, jak również istnienie dołu potencjału w odległościach $r < R$. Wykres $E_p = E_p(r)$ przedstawiony jest schematycznie na rys. 31.6. Otóż energetyczna bariera potencjalna dla cząstek α jest znacznie większa niż energia wyzwalana w czasie rozpadu α. Na podstawie mechaniki klasycznej nie można wyjaśnić przejścia cząstki α przez barierę energetyczną wyższą od energii posiadanej przez daną cząstkę.



Rys. 31.6

Rozwiązanie tego problemu w ujęciu mechaniki kwantowej podali w roku 1928 Gamow, Condon i Gurney. Obliczyli oni prawdopodobieństwo procesu przejścia w ciągu 1 sekundy cząstki α poruszającej się z prędkością v wewnątrz jądra o promieniu R przez barierę energetyczną. Otrzymali skończone, aczkolwiek na ogół nieduże, wartości prawdopodobieństwa.

W omawianym zjawisku przenikania cząstek α przez barierę energetyczną otaczającą jądro (zwanym często efektem tunelowym) występuje wyraźnie dualizm korpuskularno-falowy. Po przejściu bariery badana cząstka ma charakter korpuskularny, w czasie przenikania bariery musimy jej przypisywać naturę falową. Właśnie to założenie falowej natury cząstki α prowadzi do skończonego prawdopodobieństwa przejścia jej przez barierę energetyczną.

Wartość prawdopodobieństwa wystąpienia efektu tunelowego jest zależna od energii rozpadu α i wysokości bariery. Jeżeli są to wartości stosunkowo mało różniące się, prawdopodobieństwo przeniknięcia bariery jest dość duże, tzn. badany pierwiastek promieniotwórczy ma dużą stałą rozpadu, a mały okres połowicznego zaniku. Taki przypadek odpowiada np. $^{212}_{84}Po$, dla którego $E_\alpha = 8,8$ MeV, $\lambda = 2,31 \cdot 10^6$ s^{-1}, a $T = 3 \cdot 10^{-7}$ s. I odwrotnie, stosunkowo małej wartości E_α odpowiada małe prawdopodobieństwo wystąpienia efektu tunelowego, mała stała rozpadu, a duży okres połowicznego zaniku. Jako przykład może służyć $^{238}_{92}U$, dla którego $E_\alpha = 4,18$ MeV; $\lambda = 4,87 \times 10^{-18}$ s^{-1}; $T = 4,5 \cdot 10^9$ lat.

Przejdźmy obecnie do energii cząstek β. Zasięgi tych cząstek zbadane np. w komorze Wilsona mają różne długości, co świadczy o różnych wartościach liczbowych energii kinetycznych, a więc i prędkości początkowych, z jakimi cząstki opuszczają

jądra atomowe. Mówimy w tym przypadku o *ciągłym widmie prędkości cząstek β* rozciągającym się od wartości zerowej do pewnej maksymalnej, zbliżonej do prędkości światła w próżni. Wyjaśnienie tego faktu nastręczało wiele trudności. Jeżeli bowiem przyjąć, że wszystkie jądra danego pierwiastka w momencie emisji β mają tę samą energię i po emisji przekształcają się w jądra nowego pierwiastka również o określonej, mniejszej od początkowej energii, to każdemu aktowi emisji β musi towarzyszyć ta sama różnica energii jądra. Gdyby wylatujący elektron był jedyną cząstką emitowaną przez promieniotwórcze jądro, to stwierdzone doświadczalnie istnienie ciągłego widma prędkości (a zatem i ciągłego widma energii) przeczyłoby zasadzie zachowania energii. Zasada ta jest jednak tak ugruntowana, że należało raczej szukać innego wyjaśnienia mechanizmu przemiany β.

Wysunięto hipotezę (Pauli, r. 1931), że w przemianie β łącznie z elektronem emitowana jest z jądra dodatkowa cząstka niezmiernie przenikliwa, pozbawiona ładunku i mająca znikomą masę. Nadano tej cząstce nazwę *neutrina**.

Dalsze rozważania wykazały, że hipotetycznej na razie, nowej cząstce należy przypisać połówkowy moment pędu (tzw. spin połówkowy), aby zapewnić spełnienie zasady zachowania momentu pędu w omawianej przemianie.

Doświadczalne potwierdzenie istnienia neutrina było bardzo trudne i mimo wielokrotnie ponawianych prób dopiero w r. 1951 udało się uzyskać pierwsze przekonywające dowody. Obecnie przyjmuje się za fakt dowiedziony, że emisji β towarzyszy emisja dodatkowej, niesłychanie przenikliwej cząstki unoszącej dodatkową, nieskwantowaną część ogólnej energii związanej z daną przemianą promieniotwórczą. Dzięki emisji tej cząstki, obdarzonej spinem połówkowym, w zachodzącej przemianie jądrowej obowiązują trzy fundamentalne zasady zachowania, a mianowicie zasady zachowania energii, momentu pędu i pędu.

31.4. Detektory promieniowania

Już w poprzednich paragrafach powoływaliśmy się na różne urządzenia umożliwiające detekcję i badanie właściwości promieni wysyłanych przez ciała promieniotwórcze. W punkcie niniejszym rozszerzymy i usystematyzujemy te wiadomości.

Badanie promieniotwórczości opiera się na wykorzystaniu następujących właściwości wysyłanego promieniowania: 1) zdolności wywoływania jonizacji, 2) zdolności wywoływania świecenia pewnych substancji, 3) oddziaływania na emulsję fotograficzną, 4) stałego wydzielania ciepła przez preparaty promieniotwórcze. Omówimy przykłady detektorów, których działanie wiąże się z wyżej wymienionymi właściwościami promieniowania.

---

* Z późniejszych rozważań teoretycznych wynikło, że opisana wyżej cząstka towarzyszy obu typom przemian β, a mianowicie β^+ i β^-, przy czym równania przemian mają odpowiednio postać:

$$p \rightarrow n + e^+ + \nu,$$
$$n \rightarrow p + e^- + \tilde{\nu}.$$

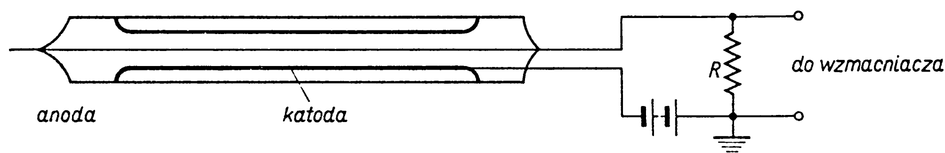
Symbole ν i $\tilde{\nu}$ oznaczają odpowiednio *neutrino* i *antyneutrino*, cząstki o znikomych masach, pozbawione ładunku elektrycznego i mające przeciwne zwroty momentów pędu.

1. Do detektorów, których działanie opiera się na zdolności wywoływania jonizacji należą m.in. komora Wilsona (por. § 30.2 (r. 1897)) i komora pęcherzykowa, pozwalające na fotografowanie torów jonizujących cząstek oraz licznik Geigera-Müllera.

Komora pęcherzykowa działa na podobnej zasadzie, co komora Wilsona, z tą jednak różnicą, że jest wypełniona nie parą przesyconą, lecz cieczą przegrzaną.

W komorze pęcherzykowej znajduje się ciecz (np. ciekły wodór, izopentan itd.) o temperaturze wyższej od normalnej temperatury wrzenia, poddana zwiększonemu ciśnieniu. W chwili gwałtownego obniżenia ciśnienia rozpoczyna się w cieczy proces wrzenia, przede wszystkim w tych miejscach, w których znajdują się jony. Działają one jako zarodki tworzenia się pary. A zatem, jeżeli zmniejszenie ciśnienia w komorze pęcherzykowej następuje w czasie przelotu przez komorę jonizującej cząstki, to jej droga zaznacza się w postaci ciągu pęcherzyków pary. Obraz taki utrzymuje się w komorze w ciągu kilku milisekund, po czym rozpoczyna się wrzenie całej cieczy.

Do zalet komór pęcherzykowych w stosunku do komór Wilsona należy zastosowanie w nich substancji o większej gęstości, a więc silniej absorbującej energię. Dzięki temu komory pęcherzykowe nadają się specjalnie do badania cząstek o wysokich energiach, wytwarzanych we współczesnych akceleratorach lub występujących w promieniach kosmicznych. Długości komór pęcherzykowych są rzędu paru metrów.



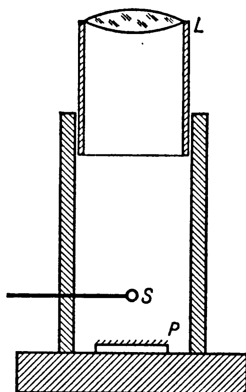
Rys. 31.7

Innym przyrządem, mającym bardzo szerokie zastosowanie, jest licznik Geigera-Müllera, pozwalający na liczenie cząstek. Budowę licznika przedstawia schematycznie rys. 31.7. Podstawową część licznika stanowi rurka metalowa lub szklana z nalotem metalu na wewnętrznej powierzchni. Wzdłuż podłużnej osi rurki rozciągnięty jest drucik. Rurka metalowa (lub wspomniany nalot przewodzący) stanowi katodę, a rozpięty drucik — anodę. Różnica potencjałów między anodą i katodą równa się około 1000 V. Każda cząstka jonizująca, przebiegająca wewnątrz licznika, wywołuje powstanie jonów. Dzięki silnemu polu między anodą i katodą powstające jony są tak rozpędzane, że same zaczynają jonizować. W ten sposób otrzymujemy bardzo znaczne wzmocnienie (amplifikację) efektu jonizacji, tak że przejściu każdej cząstki jonizującej (niezależnie od jej energii) towarzyszy powstanie silnego impulsu napięciowego na oporniku dołączonym do licznika. Impuls ten może, po dalszym wzmocnieniu, wywołać trzask w głośniku lub też może być odnotowany przez dołączony do licznika elektroniczny przelicznik impulsów.

Istnieje niebezpieczeństwo, że wobec dużej amplifikacji w przypadku trafiania do licznika wielu cząstek jonizujących licznik przestanie liczyć oddzielne impulsy, lecz stanie się trwale przewodzący. Można temu zapobiec włączając szeregowo do ob-

wodu licznika duże opory, stosując odpowiednie gaszące układy lampowe lub budując tzw. liczniki samogasnące, wypełnione specjalną mieszaniną gazów.

2. Metodę liczenia cząstek, opartą na ich zdolności wywoływania świecenia w postaci oddzielnych błysków zwanych scyntylacyjnymi, omawialiśmy już w § 30.2. Specjalnym urządzeniem służącym do liczenia scyntylacji jest tzw. *spintaryskop* (rys. 31.8). Na dnie metalowej cylindrycznej osłony znajduje się płytka *P* pokryta odpowiednią substancją, np. siarczkiem cynku. W pewnej odległości od płytki znajduje się preparat promieniotwórczy *S*. Do górnej części osłony wmontowany jest układ optyczny, np. w postaci lupy *L*, przez którą można obserwować scyntylacje, a nawet — licząc oddzielne błyski — ustalać liczbę cząstek uderzających o płytkę *P*.

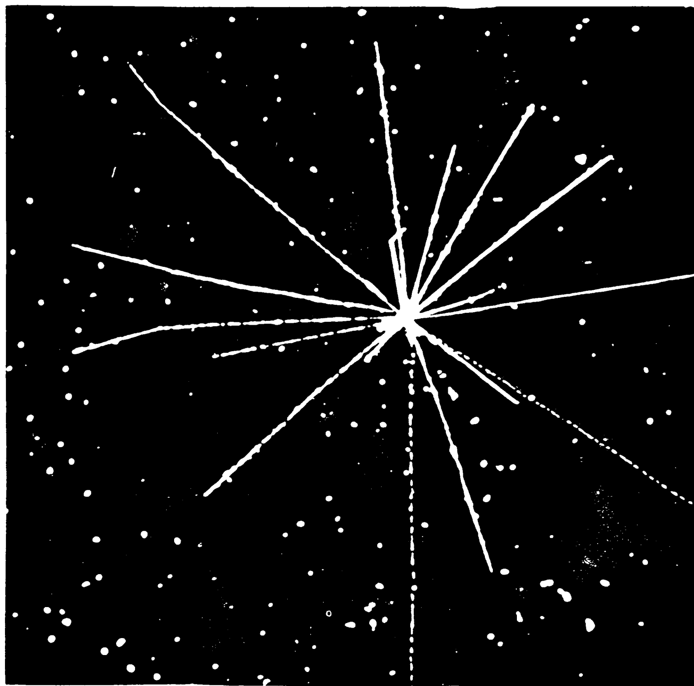


Rys. 31.8

Aczkolwiek w początkowych stadiach badań promieniotwórczości tego rodzaju urządzenia były szeroko stosowane, to jednak z biegiem czasu wobec licznych wad, związanych przede wszystkim z błędami subiektywnego pomiaru, zostały one wyparte przede wszystkim przez licznik Geigera–Müllera. Ostatnio jednak urządzenie to przeżywa renesans, dzięki sprzężeniu go z fotopowielaczem elektronowym, o którym mówiliśmy już w § 29.7. W nowoczesnych urządzeniach tego typu jako ekrany fluoryzujące stosowane są scyntylatory organiczne stałe, np. antracen, naftalen, rozpowszechniły się też scyntylatory ciekłe. Dodatkowe urządzenia elektroniczne włączone do obwodu licznika scyntylacyjnego pozwalają przeprowadzać automatycznie analizę impulsów w zależności od energii cząstek, jak również pozwalają liczyć impulsy (a więc i cząstki) odpowiadające różnym zakresom energii.

3. Do bardzo ważnych metod badania promieniotwórczości należy metoda fotograficzna. Traktujemy ją jako metodę odrębną, aczkolwiek w zasadzie opiera się ona również na zdolności jonizacji. Tym razem jonizacja zachodzi w emulsji fotograficznej (żelatyna z zawiesiną chlorku, bromku lub jodku srebra). Przechodząca przez emulsję cząstka (lub kwant promieniowania) wywołuje rozpad AgCl, AgBr lub AgI i wydzielanie się srebra tworzącego na razie niedostrzegalny dla oka tzw. obraz utajony. Proces wywoływania powoduje dalszy rozkład soli srebrowych oraz dalsze wydzielanie się

i osadzanie srebra w otoczeniu pierwotnego obrazu utajonego. Podczas procesu utrwalania reszta soli srebrowych zostaje rozpuszczona i podczas płukania usunięta z emulsji. Po wywołaniu i utrwaleniu kliszy mamy więc w emulsji zachowany obraz rozkładu substancji promieniotwórczej zawartej w badanym ciele spoczywającym przez pewien czas na kliszy. Jest to tzw. *metoda autoradiografii*, dająca informacje o przebiegu procesu promieniotwórczego w ciągu pewnego czasu.



Rys. 31.9

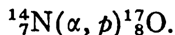
Warto tu wspomnieć o metodzie badania torów różnych cząstek naładowanych (np. pochodzących z promieni kosmicznych) za pomocą specjalnych klisz jądrowych o stosunkowo dużej grubości warstw emulsji o dużym stężeniu soli srebrowych (jodek srebra). Pod mikroskopem bada się następnie ślady dróg (rys. 31.9) poszczególnych cząstek, wyciągając z ich charakteru wnioski co do rodzaju cząstki i jej energii.

4. Do badania ciepła wytwarzanego przez substancje promieniotwórcze stosowane są specjalne mikrokalorymetry. Warto tu wspomnieć mikrokalorymetr adiabatyczny, skonstruowany przez badaczy polskich Dorabialską, Świętosławskiego i Złotowskiego, umożliwiający bardzo dokładny pomiar ciepła wytwarzanego w czasie przemian promieniotwórczych.

31.5. Reakcje jądrowe wywołane działaniem prędkich cząstek i fotonów γ

Przejdziemy z kolei do omówienia najbardziej podstawowych osiągnięć fizyków w dziedzinie przemian jądrowych, wybierając najważniejsze przykłady przemian wywołanych działaniem: a) prędkich cząstek, b) kwantów γ (reakcje fotojądrowe).

Do wywołania pierwszych sztucznych przemian pierwiastków wykorzystano cząstki α wylatujące z dużymi prędkościami z jąder pierwiastków naturalnie promieniotwórczych. W roku 1919 Rutherford stwierdził, że cząstki α przenikając przez warstwę azotu powodują przemianę typu (α, p). Ten symboliczny zapis należy rozumieć w ten sposób, że cząstką „bombardującą" jądro jest cząstka α, a wyrzuconą z jądra — jest proton p. Pełny symboliczny zapis wymienionej reakcji jądrowej ma postać



$$^{14}_{7}\mathrm{N}(\alpha, p)^{17}_{8}\mathrm{O}.$$

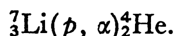
W tej pierwszej sztucznej reakcji jądrowej azot ulega przemianie na izotop tlenu o liczbie masowej 17.

Inną reakcją jądrową o podstawowym znaczeniu jest reakcja związana z odkryciem neutronu (Bothe i Becker, 1930). Można ją zapisać w postaci



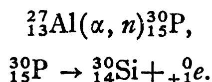
$$^{9}_{4}\mathrm{Be}(\alpha, n)^{12}_{6}\mathrm{C}.$$

Rok 1931 przynosi pierwsze zastosowanie cząstek sztucznie przyspieszanych do wywoływania reakcji jądrowych (Cockcroft i Walton oraz Lejpunski, Walter i Sinielnikow). Wywołaną reakcję można zapisać w postaci



$$^{7}_{3}\mathrm{Li}(p, \alpha)^{4}_{2}\mathrm{He}.$$

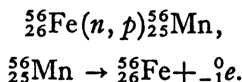
Jak już wspominaliśmy w innym miejscu (punkt 31.2.3), rok 1932 przyniósł odkrycie pozytonu przez Andersona w czasie badań promieni kosmicznych. Ta sama cząstka została zidentyfikowana przez Irenę i Fryderyka Joliot-Curie jako produkt rozpadu pierwszego, sztucznie otrzymanego w reakcji (α, n), pierwiastka promieniotwórczego $^{30}_{15}\mathrm{P}$. Kolejne reakcje przedstawiają się następująco:



$$^{27}_{13}\mathrm{Al}(\alpha, n)^{30}_{15}\mathrm{P},$$

$$^{30}_{15}\mathrm{P} \rightarrow {}^{30}_{14}\mathrm{Si} + {}^{\;\,0}_{+1}e.$$

Innym typem reakcji jest przemiana jądrowa wywołana bombardowaniem neutronami. Ten rodzaj bardzo skutecznych pocisków jądrowych zastosował po raz pierwszy w swych badaniach Fermi w 1934 roku. Przykładem takiej reakcji może być

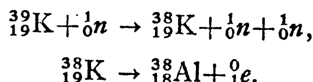


$$^{56}_{26}\mathrm{Fe}(n, p)^{56}_{25}\mathrm{Mn},$$

$$^{56}_{25}\mathrm{Mn} \rightarrow {}^{56}_{26}\mathrm{Fe} + {}^{\;\,0}_{-1}e.$$

Do wywoływania tego rodzaju przemian nadają się szczególnie neutrony powolne. Można je otrzymać przez przepuszczenie neutronów prędkich przez warstwy substancji bogatych w wodór, np. przez parafinę, wodę itp. Wobec bardzo zbliżonych wartości mas atomów wodoru i neutronów zachodzą zderzenia sprężyste (por. § 8.8), podczas których neutrony znacznie zmniejszają swój pęd. Następnie po znalezieniu się w pobliżu jądra takie powolne neutrony zostają do niego wciągnięte potężnymi siłami jądrowymi. Nabywają więc dużej energii kinetycznej, którą przekazują innym cząstkom w jądrze. W wyniku takiego „wtargnięcia" nowej cząstki do jądra może powstać — jak widzieliśmy z przytoczonych wyżej przykładów — nowe, nietrwałe jądro ulegające rozpadowi (promieniotwórczość sztuczna) i może też z jądra wylecieć inny element jego budowy, który uzyskał dodatkową energię.

Omówiony wyżej przykład nie jest jedynym typowym dla reakcji zachodzących pod wpływem neutronów. Oprócz reakcji typu (n, p) i (n, α) znane są reakcje (n, γ), podczas których również wytwarza się radioaktywny izotop ulegający rozpadowi z emisją elektronu, jak również typu $(n, 2n)$ z dalszym rozpadem promieniotwórczym z emisją pozytonu. Na przykład



$$^{39}_{19}K + ^{1}_{0}n \rightarrow ^{38}_{19}K + ^{1}_{0}n + ^{1}_{0}n,$$

$$^{38}_{19}K \rightarrow ^{38}_{18}Al + ^{0}_{1}e.$$

Nie będziemy dalej mnożyć przykładów. Ograniczymy się tylko do podania niepełnego zresztą zestawienia typów reakcji jądrowych, w celu zilustrowania różnorodności przemian wywołanych działaniem cząstek prędkich.

Typy reakcji jądrowych

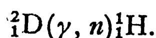
(α, α)	(α, p)	(α, n)	$(\alpha, 2n)$	(α, γ)	
(d, α)	(d, p)	(d, n)	$(d, 2n)$		
(p, α)	(p, d)	(p, p)	(p, n)	$(p, 2n)$	(p, γ)
(n, α)	(n, p)	(n, n)	$(n, 2n)$	(n, γ)	

Symbol d w podanym zestawieniu oznacza deuteron—jądro ciężkiego wodoru.

Ostatnio do wywołania reakcji jądrowych zaczęto stosować przyspieszone ciężkie jony, np. $^{12}_{6}C^{6+}$ lub $^{14}_{7}N^{6+}$.

O reakcjach rozszczepienia jąder ciężkich na dwa fragmenty była już mowa w punkcie 31.2.3. Do tego zagadnienia wrócimy jeszcze w § 31.7. Warto podkreślić, że prócz reakcji rozszczepienia znane są reakcje kruszenia, podczas których bombardowane jądro rozpada się na wiele fragmentów.

Odrębny rodzaj reakcji jądrowych stanowią tzw. reakcje fotojądrowe (wywołane działaniem fotonów γ). Jako przykład może służyć reakcja o zapisie symbolicznym (γ, n), a więc wywołana działaniem promieniowania γ i połączona z emisją np. neutronu. Taki przebieg ma reakcja fotojądrowa, odkryta w r. 1934 przez Chadwicka i Goldhabera



$$^{2}_{1}D(\gamma, n)^{1}_{1}H.$$

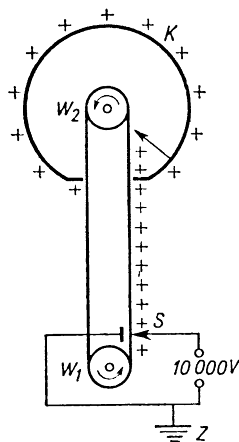
Na podstawie tej reakcji wyznaczono masę neutronu. Istnieją też reakcje fotojądrowe typu $(\gamma, p+n)$, $(\gamma, \alpha+n)$ itp.

31.6. Akceleratory

W § 31.5 omawialiśmy różne reakcje jądrowe wywołane działaniem prędkich cząstek otrzymanych czy to w czasie naturalnych przemian promieniotwórczych, czy też dzięki przyspieszeniu ich w specjalnych urządzeniach, zwanych ogólnie akceleratorami. Rozpatrzmy te sprawy dokładniej.

Wydajność procesu dezintegracji przy użyciu promieni α lub β z naturalnych źródeł promieniotwórczych była mała m.in. dlatego, że nie można było uzyskać strumieni cząstek celnie skierowanych. Nic więc dziwnego, że szukano nowych źródeł cząstek o dużych energiach — źródeł bardziej wydajnych i szerzej dostępnych.

Przede wszystkim wykorzystano źródła wysokich napięć do stworzenia pól elektrycznych rozpędzających cząstki naładowane. Nadają się do tego celu transformatory wysokiego napięcia użyte łącznie z lampami prostowniczymi (kenotronami — por. lampy rentgenowskie). Dużą trudność stanowiło osiągnięcie dobrej izolacji zwojów, zapobiegającej iskrzeniu wewnątrz transformatorów i na elektrodach kenotronów. Wyprostowane, wysokie napięcie można już wykorzystywać do rozpędzania jonów w tzw. akceleratorach. Najczęściej stosowanymi cząstkami są protony, deuterony i jądra helu (cząstki α). Otrzymuje się je w rurach próżniowych akceleratora. Rury te są zwykle wykonane z porcelany, ich średnice osiągają kilkadziesiąt centymetrów, a długości przewyższają 1 metr. W rurze istnieją dwie oddzielne komory: w pierwszej pod wyższym ciśnieniem, rzędu setnych milimetra Hg, wytwarzane są jony, w drugiej — działaniem pomp utrzymuje się możliwie najlepszą próżnię i rozpędza jony. Rozpędzanie odbywa się stopniowo, dzięki zastosowaniu szeregu elektrod o rosnących potencjałach. Zastosowanie szeregu elektrod zapewnia równocześnie „soczewkowe", skupiające działanie pola na strumień jonów. Na końcu rury znajduje się cel, tzn. materiał poddany bombardowaniu, zwany często tarczą. Cząstki emitowane przez jądra atomowe tarczy podczas bombardowania (a nawet po bombardowaniu w przypadku sztucznej promieniotwórczości) są dalej badane w odpowiednich detektorach (komora Wilsona, liczniki Geigera itd.)



Rys. 31.10

Szerokie zastosowanie w fizyce jądrowej znalazł *akcelerator elektrostatyczny*, zwany *generatorem van de Graaffa*, za pomocą którego można osiągać napięcie rzędu kilku milionów woltów.

Działanie generatora van de Graaffa opiera się na zjawisku indukcji elektrostatycznej oraz na właściwości przewodników metalicznych gromadzenia ładunków wyłącznie na powierzchni zewnętrznej.

Zasadę działania generatora van de Graaffa przedstawia rys. 31.10.

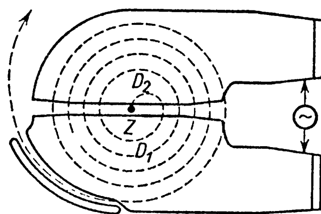
W wysokiej wieży, między dwoma wałami W_1 i W_2, przebiega pionowo jedwabny pas. Wał górny W_2 znajduje się wewnątrz wielkiej wydrążonej metalowej kuli K o średnicy około 4 m, osadzonej na wieży. W dolnej części wieży znajduje się źródło wysokiego, stałego napięcia rzędu 10^4 V. Z tego źródła, za pośrednictwem ostrza S, ładunki np. dodatnie spływają na pas, który w swoim ruchu (wywołanym ruchem obrotowym wałów) przenosi je do wnętrza kuli. Tam — dzięki układowi kolców wychodzących z wewnętrznej powierzchni kuli — ładunki ściągane są z pasa na kulę i grupują się na jej zewnętrznej powierzchni, a część pasa pozbawiona już ładunków przesuwa się ku dołowi po nową porcję ładunków ze źródła.

Analogiczny proces odbywa się w drugiej takiej samej wieży z tą tylko różnicą, że na powierzchnię drugiej kuli dostarczane są ładunki ujemne. W ten sposób można osiągnąć między kulami napięcie rzędu kilku megawoltów (1 MV = 10^6 V). Napięcie to można już wykorzystywać do przyspieszania jonów umieszczając między kulami odpowiednie rury próżniowe i kierując przez nie strumienie jonów. Generator znalazł

szerokie zastosowanie w badaniach jądrowych, jak również przy wytwarzaniu bardzo przenikliwych promieni Röntgena.

Innym typem akceleratora jest zbudowany przez Lawrence'a tzw. *cyklotron*, akcelerator należący do grupy *akceleratorów magnetycznych*. W urządzeniu tym wyeliminowana została trudność wytwarzania bardzo silnych pól elektrycznych przyspieszających, zastosowano bowiem przyspieszanie wielokrotne podczas powtarzającego się przebiegu badanej cząstki przez niezbyt silne elektryczne pole przyspieszające. Ten rodzaj akceleratora przypomina kształtem płaskie cylindryczne pudło metalowe (rys. 31.11), rozcięte wzdłuż średnicy, o nieco rozsuniętych obu połówkach, zwanych duantami. Wewnątrz obu duantów utrzymywana jest wysoka próżnia. W pobliżu środka szczeliny między duantami umieszczone jest źródło jonów Z. Całość znajduje się między nabiegunnikami silnego elektromagnesu. Pole magnetyczne prostopadłe do płaszczyzny rysunku zmusza jony do krążenia po torach zakrzywionych. Między duantami przyłożone jest zmienne napięcie o wysokości kilkudziesięciu tysięcy woltów. Częstotliwość zmian napięcia jest tak dobrana, aby pole elektryczne działało przyspieszająco na jon przy każdym jego przejściu przez szczelinę między duantami. Wewnątrz duantów pola elektrycznego nie ma, tam więc odbywa się ruch jednostajny po półokręgu pod wpływem pola magnetycznego. Łatwo można sprawdzić, od jakich czynników zależy czas obiegu jednego półokręgu. Niech jon o ładunku np. $+Q$ i masie m po przejściu szczeliny między duantami ma prędkość v. Czas t przelotu jonu wzdłuż półokręgu o promieniu r wewnątrz duantu równa się

$$t = \frac{\pi r}{v}.$$



Rys. 31.11

Ruch jonu odbywa się pod działaniem pola magnetycznego o indukcji B, a zatem

$$\frac{mv^2}{r} = QvB. \tag{31.3}$$

Z układu dwóch ostatnich równań otrzymujemy

$$t = \frac{\pi m}{QB}. \tag{31.4}$$

Wynik jest interesujący. Czas przelotu jonu wzdłuż półokręgu nie zależy ani od prędkości v jonu, ani od promienia r. Można zatem tak dobrać urządzenie wytwarzające napięcie między duantami (tzn. generator lampowy dostarczający szybkozmiennego

napięcia), by półokres zmian napięcia był równy czasowi przelotu jonu wzdłuż półokręgu. Zobaczmy, jakie korzyści daje spełnienie tego warunku. Rozważmy jeden pełny obieg jonu. Po przejściu przez szczelinę pod działaniem napięcia przyspieszającego w kierunku od D_1 do D_2 jon obiega półokrąg w duancie D_2 z prędkością v. Po upływie czasu t dostaje się on ponownie do przerwy między duantami zastając tam zmienione co do znaku napięcie. Tym razem pole elektryczne przyspiesza ruch jonu w stronę duantu D_1. Z nieco większą prędkością v' a więc i po kole o nieco większym promieniu, jak to widać ze wzoru (31.3), zaczyna się obieg półokręgu w duancie D_1. Znowu po upływie czasu t jon przebiega przestrzeń między duantami w kierunku od D_1 do D_2 zwiększając swą prędkość z v' na v'' itd. Każdemu przejściu między duantami towarzyszy wzrost prędkości ruchu i powiększenie promienia orbity. Po wielokrotnie powtórzonych okrążeniach (ruch po spirali), gdy już rozpędzona cząstka bliska jest obwodu „puszki" cyklotronu, zostaje ona za pomocą pomocniczego pola odchylona w bok od właściwego dla niej toru kołowego i wyprowadzona na zewnątrz przez cienkie okienko. W niewielkiej odległości od okienka znajduje się tarcza, czyli substancja przeznaczona do bombardowania. Wyniki bombardowania badane są w komorach lub licznikach. Bombardowanie badanej substancji może powodować sztuczną promieniotwórczość, produkcję neutronów itp.

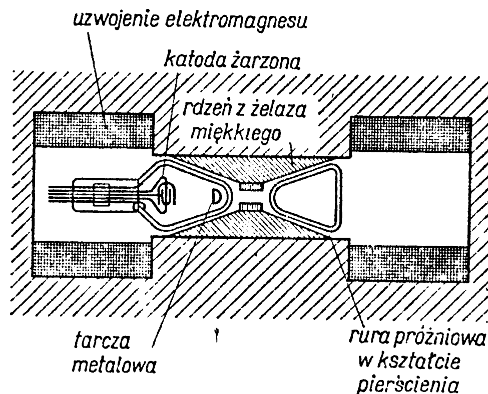
Wróćmy jednak jeszcze do źródła jonów Z, o którym krótko powiedzieliśmy, że znajduje się w pobliżu środka szczeliny między duantami. Ściślej mówiąc, mamy tam żarzony drucik stanowiący źródło elektronów. Elektrony, jako cząstki o małej masie podlegając działaniu pola magnetycznego, zaczynają krążyć po torach o małych promieniach krzywizny, tzn. w pobliżu środka szczeliny. W tym właśnie obszarze na skutek zderzeń z nielicznymi atomami gazu wywołują one jonizację. Wytworzone jony (protony, deuterony, jądra helu) rozpoczynają od razu krążenie po wyżej omówionych torach spiralnych.

Napięcia stosowane w cyklotronie mają wartości w granicach 10–200 kV. Energia wylatujących cząstek może osiągać wartość około 25 MeV dla protonów, a dla cząstek α i deuteronów jeszcze większą. Można więc cyklotron uważać za niskonapięciowe źródło cząstek o dużej energii.

Jak już mówiliśmy, rozpędzone w cyklotronie jony można wykorzystać do produkcji neutronów. Trzeba jednak pamiętać, że wiąże się z tym poważne niebezpieczeństwo dla obsługi cyklotronu i ludzi znajdujących się w bliskim otoczeniu. Promieniowanie neutronowe jest niesłychanie przenikliwe, może przenikać np. przez ściany budynków produkując na swej drodze pierwiastki sztucznie promieniotwórcze. W organizmie ludzkim zachodzą również przemiany promieniotwórcze prowadząc przede wszystkim do tzw. anemii złośliwej. Dobrą osłonę aparatury produkującej neutrony stanowi np. warstwa wody o grubości 1 metra, która zmniejsza liczbę neutronów do 0,01%.

Cyklotron nie nadaje się do przyspieszania elektronów, gdyż wobec ich małej masy już przy energii bliskiej 1 MeV mają one prędkości tak wielkie, że występują relatywistyczne zmiany masy. Zmienność masy odbija się na czasie przelotu wzdłuż półokręgu (por. 31.4) i powoduje zanik synchronizacji między obiegiem półokręgu i zmianami znaku pola elektrycznego między duantami.

W roku 1941 Kerst zbudował pierwsze urządzenie przeznaczone specjalnie do nadawania dużych energii elektronom, tzw. *betatron*, zaliczany do *akceleratorów indukcyjnych*. W betatronie otrzymuje się wiązki elektronowe o dużej energii (obecnie już rzędu setek MeV). Wiązki te skierowane na tarcze powodują emisję promieniowania γ o energii tego samego rzędu. Otrzymane w ten sposób fotony γ wykorzystywane są do wywoływania sztucznych przemian fotojądrowych. Zasadę budowy i działania betatronu przedstawia rys. 31.12. Źródłem elektronów jest żarzone włókno stanowiące katodę. Dzięki zastosowaniu odpowiedniego pola elektrycznego (rozmieszczenie anody i siatki w pobliżu katody) uzyskuje się zogniskowane wiązki elektronów skierowane do wnętrza komory w kształcie pierścienia. W komorze mamy wysoką próżnię, utrzymywaną dzięki stałemu działaniu pomp. Całość urządzenia znajduje się między nabiegunnikami elektromagnesu wytwarzającego zmienne pole magnetyczne o kierunku równoległym do osi urządzenia. Do dobrego funkcjonowania przyrządu potrzebne jest pole magnetyczne niejednorodne w obszarze całego pola orbity. W wyniku kilkunastoletnich prób udało się przez dobór kształtu nabiegunników elektromagnesu otrzymać odpowiednie pole magnetyczne. Działanie tego pola zapewnia: 1) krążenie elektronu po orbicie kołowej o stałym promieniu pod działaniem siły evB w płaszczyźnie prostopadłej do płaszczyzny rysunku oraz 2) przyspieszenie elektronów na orbicie dzięki powstawaniu siły elektromotorycznej w płaszczyźnie toru elektronu. (Tor elektronu można traktować jako odpowiednik obwodu prądu elektrycznego.) Strumień indukcji magnetycznej jest tak dobrany, że elektrony poruszają się po orbicie o stałym promieniu.



Rys. 31.12

Elektrony dokonują setek tysięcy obiegów w ciągu $\frac{1}{4}$ okresu zmienności pola magnetycznego (przejście od wartości zerowej pola do maksimum). W czasie każdego obiegu następuje pewien przyrost energii elektronów. Gdy ich energia osiąga zamierzoną wartość, wywołane zostaje określone zaburzenie w strumieniu indukcji magnetycznej, które zmienia promień orbity i skierowuje wiązkę elektronów na tarczę. W nowoczesnych betatronach udało się otrzymać energię około 1000 MeV. Jest to równocześnie maksymalna wartość energii fotonów powstających podczas bombardowania

tarczy elektronami. Jak już mówiliśmy, fotony te mogą być inicjatorem reakcji jądrowych, np. typu (γ, n), $(\gamma, p+n)$, $(\gamma, \alpha+n)$.

Nowszego typu cyklotrony (tzw. *synchrocyklotrony*, *synchrotrony*, *synchrofazotrony*, zwane też *synchrotronami protonowymi*), eliminujące wpływ relatywistycznych zmian masy, budowane są w różnych krajach od r. 1945 do chwili obecnej. Jest to swego rodzaju wyścig międzynarodowy w kierunku wprowadzania udoskonaleń i osiągania wyższych energii. Trzeba tu podkreślić wielkie osiągnięcia, jakie mają ZSRR (Dubna i Sierpuchow), USA (Brookhaven) oraz Europejski Ośrodek Badań Jądrowych w Genewie. Faktem jest już nadawanie protonom energii dochodzącej do 100 gigaelektronowoltów (GeV).

31.7. Reakcje rozszczepienia jąder atomowych. Model kroplowy jądra

Lata 1938–1939 przyniosły w dziedzinie fizyki jądrowej nowe odkrycia o niesłychanie ważnych konsekwencjach. Hahn i Strassmann stwierdzili, że podczas bombardowania neutronami jąder uranu pojawia się pierwiastek bar o liczbie atomowej $Z = 56$. (To samo stwierdzono następnie w przypadku toru.) Innymi słowy, z pierwiastka ciężkiego o liczbie Z bliskiej wartości 90 otrzymuje się pierwiastek należący do środkowej części układu okresowego. Dosłownie w kilka dni później Frisch i Liza Meitner podali poprawną interpretację tego zjawiska. Ciężkie jądro po wchłonięciu neutronu staje się jądrem nietrwałym i ulega rozszczepieniu na dwa fragmenty o zbliżonej liczbie masowej. Powstają dwa jądra pierwiastków ze środkowej części układu okresowego. Dalsze badania wykazały, że wynik rozszczepienia może być różny. Jednym z przykładów reakcji rozszczepienia może być:



$$^{235}_{92}\mathrm{U}+{}^{1}_{0}n \to {}^{236}_{92}\mathrm{U} \to {}^{94}_{38}\mathrm{Sr}+{}^{140}_{54}\mathrm{Xe}+2\,{}^{1}_{0}n+\mathrm{Q}. \qquad (31.5)$$

Odkrycie reakcji rozszczepienia jąder ciężkich miało nie tylko olbrzymie znaczenie poznawcze, lecz i praktyczne. Była to pierwsza reakcja, w której liczba atomowa wyjściowego jądra atomowego zmieniała się więcej niż o 2 jednostki i w której produkt wyzwalany był cięższy od cząstki α.

Z punktu widzenia praktycznego znaczenie odkrytej reakcji polega na wyzwalaniu olbrzymiej ilości energii. Łatwo sprawdzić korzystając z rys. 31.1, że energia wiązania jednego nukleonu w jądrze uranu jest o 1 MeV mniejsza od tejże energii w jądrach pierwiastków znajdujących się blisko środka układu okresowego. A zatem w procesie przekształcenia się 1 jądra uranu w 2 różne jądra o liczbie Z bliskiej 45 wyzwala się około 200 MeV energii.

Dla podkreślenia olbrzymiej wartości tej energii podajemy zestawienie energii różnych cząstek (tabela 31.2).

Reakcje rozszczepienia znalazły uzasadnienie teoretyczne. Frenkiel, a następnie Bohr i Wheeler wykorzystali do tego celu tzw. model kroplowy jądra podany przez Gamowa (1930), a następnie rozwinięty przez Bohra. Poświęcimy mu nieco uwagi, gdyż na podstawie tego modelu można wytłumaczyć szereg zjawisk jądrowych.

Tabela 31.2

Wartości energii różnych cząstek

Rodzaj cząstki	E	
	J	eV
cząsteczka gazu w temperaturze pokojowej	$4 \cdot 10^{-21}$	0,025
foton widzialny	około $3{,}2 \cdot 10^{-19}$	2,0
foton nadfioletowy	$4{,}8 \cdot 10^{-19}$–$1{,}6 \cdot 10^{-17}$	3–100
twardy foton rentgenowski	$1{,}6 \cdot 10^{-14}$–$1{,}6 \cdot 10^{-13}$	10^5–10^6
foton γ	$1{,}6 \cdot 10^{-13}$–$4{,}8 \cdot 10^{-13}$	10^6–$3 \cdot 10^6$
cząstka α emitowana przez rad	$7{,}36 \cdot 10^{-13}$	$4{,}8 \cdot 10^6$
cząstki promieniowania kosmicznego	$1{,}6 \cdot 10^{-10}$–$1{,}6 \cdot 10^{-9}$	10^9–10^{10}

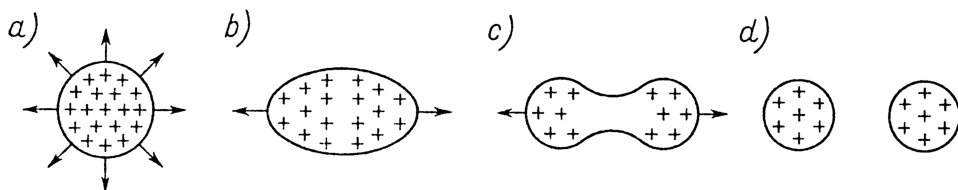
Model kroplowy jądra opiera się na analogii między siłami międzycząsteczkowymi w cieczach (zwanymi siłami van der Waalsa) i siłami jądrowymi. Zarówno pierwsze, jak i drugie siły są siłami małego zasięgu. Promień sfery działania cząsteczki cieczy jest rzędu 10^{-7} cm, a więc tego samego rzędu co promień cząsteczki. W związku z tym dana cząsteczka związana jest siłami van der Waalsa tylko ze swymi najbliższymi „sąsiadkami". Podobna sytuacja istnieje w jądrze atomowym: silne oddziaływania jądrowe występują w odległości rzędu 10^{-15} m od środka nukleonu, a więc w odległości rzędu promienia nukleonu. Efekt znów jest taki, że w grę wchodzą tylko oddziaływania przyciągające między najbliższymi nukleonami. Zarówno w cieczach, jak i w jądrze atomowym takie oddziaływania dają w wyniku gęste upakowanie elementów składowych, decydujące w zasadzie o gęstości odpowiedniej cieczy. „Ciecz jądrowa" wskutek silnego upakowania nukleonów ma olbrzymią średnią gęstość (w przybliżeniu jednakową dla wszystkich jąder atomów), rzędu 10^{17} kg/m^3, tzn. 1 cm^3 cieczy jądrowej ma masę około 150 miliardów kilogramów. Dla porównania podamy, że w warunkach ziemskich maksymalną gęstość równą 22,8 kg/m^3 ma iryd (1 cm^3 irydu ma masę 0,0228 kg).

W jądrach atomowych, podobnie jak w cieczach, istnieje dążność do osiągania stanu minimum energii potencjalnej. Stanowi temu odpowiada minimum powierzchni swobodnej. Ciecz jądrowa, podobnie jak zwykłe ciecze, wykazuje zatem objawy napięcia powierzchniowego. Każde jądro przyjmuje kształt kulistej kropli, gdyż, jak wiadomo, kula jest bryłą, którą przy danej objętości charakteryzuje minimalna powierzchnia.

Omówimy szerzej analogię między kroplą zwykłej cieczy i kroplą cieczy jądrowej. Wiadomo, że ogrzanie cieczy zwiększa energię kinetyczną poszczególnych cząsteczek i może doprowadzić do oderwania się cząsteczki od reszty cieczy, czyli do procesu parowania. W odniesieniu do cieczy jądrowej proces parowania sprowadza się do wyzwolenia z jądra nukleonu. Jak już wiemy, jest to możliwe po doprowadzeniu porcji energii, np. przez bombardowanie szybkimi cząstkami. Wlatująca do jądra cząstka α, proton lub neutron, przekazuje składnikom jądra swą energię, co wpływa na wzrost średniej energii kinetycznej nukleonów. W ujęciu teorii kinetycznej wzrostowi energii kinetycznej towarzyszy wzrost temperatury. Gdy przekazana jądru energia jest rzędu 8 MeV, to jak wynika z obliczenia, temperatura jądra wzrasta do kilku miliardów stopni. (Dla porównania podamy, że najwyższe temperatury wnętrza gwiazd szacowa-

ne są na miliony stopni.) Tę część zjawiska można porównywać do zjawiska skraplania. Skraplanie pary zwykłej cieczy związane jest z wydzielaniem ciepła w postaci emisji w podczerwieni. „Skraplanie" odpowiadające dołączeniu do jądra dodatkowego nukleonu prowadzi, jak widzieliśmy, do olbrzymiego wzrostu temperatury i emisji promieniowania γ o dużej energii. Energia oddana na zewnątrz jest równa defektowi masy jądra, a równocześnie stanowi energię wiązania nukleonu w jądrze. Czasem po emisji γ jądro ma jeszcze nadmiar energii, którego pozbywa się przez emisję β^+ lub β^- wraz z odpowiednią odmianą neutrina. Możliwy jest także inny proces jądrowy: wlatująca cząstka może w sprzyjających warunkach przekazać swą energię nie zespołowi wszystkich nukleonów w jądrze, lecz jednemu z nich (względnie zespołowi $2n+2p$). W tym przypadku „wyparowuje" z jądra właśnie ten „obdarowany" energią składnik jądra, zachodzi reakcja jądrowa.

Interesujący jest przypadek, gdy badane jądro wychwytuje szybki neutron o energii kinetycznej $E_k > 8$ MeV, tzn. większej od przeciętnej energii wiązania neutronu w jądrze. W tych warunkach neutron osiągając „dół potencjału" we wnętrzu jądra ma zasób energii wystarczający na skompensowanie energii wiązania dwóch neutronów wchodzących w skład jądra bombardowanego. W warunkach takiej przemiany rośnie liczba pocisków neutronowych i reakcja jądrowa może rozwijać się łańcuchowo (lawinowo). Omówimy to szerzej w § 31.8.



Rys. 31.13

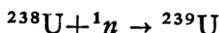
Prześledźmy jeszcze inną ewentualność związaną z wychwytem neutronu przez ciężkie jądro atomowe. Niech tym razem energia neutronu (złożona oczywiście z jego energii kinetycznej, posiadanej przed osiągnięciem sfery działania jądra i energii uzyskanej dzięki „spadkowi" do „dołu potencjału") zostaje zużyta na zwiększenie amplitud drgających nukleonów. Pociągnie to za sobą deformację kropli jądrowej (schematyczny rys. 31.13). Podczas powiększania się tej deformacji może wystąpić przewężenie kropli, a nawet jej rozerwanie na dwie oddzielne krople jądrowe. W schemacie tym odnajdujemy wyjaśnienie procesu rozszczepienia jądra.

Gdyby kropla jądrowa pozbyła się wcześniej nadmiaru energii (np. przez emisję γ), to wróciłaby bez rozszczepienia do stanu pierwotnego.

Z tego obrazowego opisu widać, że do wywołania zjawiska rozszczepienia potrzebne jest doprowadzenie pewnej początkowej porcji energii, która wywołałaby odkształcenie kropli. Dalsze badania wykazały, że wartość tej *energii aktywacji* zależy od tzw. *parametru rozszczepienia* równego Z^2/A. Im większy jest parametr rozszczepienia, tym mniejsza jest energia aktywacji, czyli tym łatwiej zachodzi rozszczepienie. Różnice

energii aktywacji mogą być dość znaczne, nawet w przypadku izotopów tego samego pierwiastka. Na przykład dla uranu ^{238}U, najbardziej rozpowszechnionego w przyrodzie, $E_{akt} = 7$ MeV, a dla izotopu ^{235}U energia $E_{akt} = 6{,}6$ MeV. Wiążą się z tym ważne konsekwencje. Przeanalizujmy możliwości wywołania rozszczepienia obu tych izotopów uranu przez wychwyt neutronów.

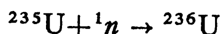
W przypadku ^{238}U reakcja wychwytu neutronu



$$^{238}U + {}^{1}n \rightarrow {}^{239}U$$

dostarcza jądru tylko 6 MeV energii dzięki „spadkowi" neutronu do „dołu potencjału". Ponieważ potrzebna energia aktywacji wynosi 7 MeV, brakująca energia, równa 1 MeV, musi być dostarczona w postaci energii kinetycznej neutronu. A zatem reakcję rozszczepienia wywołają tym razem tylko neutrony prędkie.

W przypadku ^{235}U reakcja wychwytu neutronu



$$^{235}U + {}^{1}n \rightarrow {}^{236}U$$

dostarcza jądru energii 6,8 MeV dzięki „spadkowi" neutronu do „dołu potencjału". Ponieważ tym razem energia aktywacji wynosi 6,6 MeV, więc nawet najpowolniejsze (tzw. *termiczne*) *neutrony* powodują reakcję rozszczepienia.

Może warto podkreślić, że ze względu na istnienie efektu tunelowego możliwe jest rozszczepienie w przypadku dostarczenia jądru energii mniejszej od energii aktywacji, a nawet rozszczepienie samorzutne. Prawdopodobieństwa takich procesów są jednak znacznie mniejsze. Na przykład szacuje się, że na 1800 rozpadów ^{238}U z emisją α przypada jedno rozszczepienie samorzutne.

31.8. Reakcje łańcuchowe

Już w poprzednim paragrafie wspominaliśmy, że reakcje rozszczepienia jąder uranu mogą mieć różny przebieg. Podaliśmy przykładowo symboliczny zapis (31.5) jednej z możliwych reakcji. Okazuje się, że w czasie takich reakcji jednemu wychwytowi neutronu, powodującemu rozszczepienie, towarzyszy wyzwolenie średnio 2–3 nowych neutronów. Nawet wyzwalanie jednego tylko nowego neutronu, jeśli podlegałby on również wychwytowi i wywoływał następne rozszczepienie — wystarczyłoby do tego, aby reakcja sama się podtrzymywała, tzn. żeby miała charakter *reakcji łańcuchowej*. Jeśli podczas jednego rozszczepienia wyzwala się więcej niż jeden neutron, to reakcja może rozwijać się nie tylko łańcuchowo, lecz wręcz lawinowo, wyzwalając w każdej następnej „generacji" coraz więcej neutronów. Tak (rys. 31.14) byłoby w przypadku idealnym, gdyby każdy nowo wyzwolony neutron powodował kolejne rozszczepienie. Przeanalizujmy to zagadnienie nieco dokładniej, ze względu na jego znaczenie dla możliwości wykorzystania energii jądrowej.

Zastanówmy się przede wszystkim nad dalszym losem neutronów wyzwalanych podczas reakcji rozszczepienia uranu. Może nastąpić:

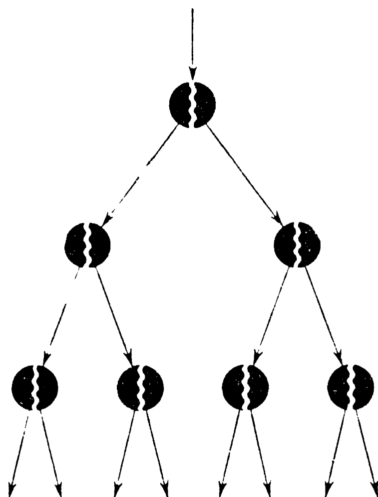
a) zderzenie sprężyste z jądrem uranu (zachodzi przy tym nie mająca istotnego znaczenia zmiana kierunku ruchu neutronu, prawie bez zmiany wartości prędkości **v**),

b) zderzenie niesprężyste z jądrem uranu (— przekazanie jądru uranu energii neutronu, powodujące tylko „ogrzanie" jądra),

c) wychwyt neutronu przez jądro uranu, prowadzący do powstania nowego izotopu, bez wywołania rozszczepienia,

d) wychwyt neutronu przez jądro uranu, prowadzący do powstania nowego izotopu, ulegającego rozszczepieniu na dwa fragmenty, z wyzwoleniem nowych neutronów,

e) ucieczka neutronu poza obszar wypełniony materiałem rozszczepialnym. Każde z wymienionych zdarzeń może mieć miejsce i każde zachodzi z pewnym prawdopodobieństwem. A zatem: 1) przebieg rzeczywistej reakcji odbiega od idealnego, przy którym wszystkie wyzwalane neutrony powinny powodować zdarzenie d, 2) reakcja rzeczywista przebiega łańcuchowo (lub nawet lawinowo), jeśli prawdopodobieństwo a i d przeważa nad wszystkimi pozostałymi. Wpływ rozmiarów obszaru wypełnionego materiałem rozszczepialnym omówimy oddzielnie. Na razie ograniczymy się do analizy zdarzeń a, b, c, d w przypadku uranu.



Rys. 31.14

Okazuje się, że różne izotopy uranu wykazują różne prawdopodobieństwa wymienionych wyżej zdarzeń, a mianowicie: w izotopie ^{235}U istnieje przy wszystkich prędkościach neutronu wybitna przewaga zdarzeń a i d nad pozostałymi. Najbardziej wydajne są neutrony termiczne. (Średnia prędkość termiczna neutronu w temperaturze 0°C wynosi 2500 m/s.) W izotopie uranu ^{238}U największe prawdopodobieństwo odpowiada zdarzeniu b. Towarzyszy temu spowolnienie neutronów. W pewnym przedziale prędkości, w tzw. *przedziale rezonansowym* (prędkości bliskie 8500 m/s) istnieje duże prawdopodobieństwo wychwytu nie prowadzącego do rozszczepienia (a więc zdarzenia c). Rozszczepienie (czyli zdarzenie d występuje tylko przy dużych prędkościach neutronów i to z małą wydajnością. Zachowania się izotopu ^{234}U nawet nie omawiamy, ze względu na znikomą jego zawartość w uranie naturalnym*.

* Przypominamy skład procentowy uranu naturalnego: ^{238}U — 99,2739%, ^{235}U — 0,7204%, ^{234}U — 0,0057%.

Jak widać z tego zestawienia, w czystym uranie ^{238}U nie wystąpi reakcja łańcuchowa. Podobnie w próbkach naturalnego uranu z olbrzymią przewagą ^{238}U nie rozwinie się reakcja łańcuchowa, jeśli nie stworzymy specjalnych warunków (por. następny paragraf). Mogłaby ona natomiast mieć miejsce w czystej bryle z uranu ^{235}U. Wybuchowi takiej bryły ^{235}U można by zapobiec wykorzystując pomijaną w dotychczasowych rozważaniach możliwość ucieczki neutronów z obszaru bryły. W tym celu należałoby próbkom z ^{235}U nadawać rozmiary tak niewielkie, aby 60% wyzwalanych neutronów wydostawało się poza próbkę, bez wywołania rozszczepienia. W tych warunkach mówi się o rozmiarach próbek mniejszych od krytycznych (*podkrytycznych*). W takich próbkach liczba neutronów maleje wykładniczo z biegiem czasu.

Gdy rozmiary próbki są tak dobrane, że liczba neutronów wewnątrz jest stała w czasie, próbka ma *rozmiary krytyczne*. Jeśli natomiast z biegiem czasu liczba neutronów w próbce rośnie wykładniczo, próbka ma *rozmiary* większe od krytycznych (*nadkrytyczne*). Zmiana rozmiarów próbki ma wpływ na liczbę neutronów w jej wnętrzu, gdyż w danym czasie liczba wyzwalanych neutronów rośnie proporcjonalnie do objętości (w przypadku kuli — do R^3), a liczba uciekających neutronów — proporcjonalnie do powierzchni (a więc w przypadku kuli — do R^2). A zatem zwiększając objętość wywołujemy wzrost liczby istniejących w próbce neutronów, czyli przesuwamy szalę na stronę przebiegu reakcji łańcuchowej (lawinowej), a zmniejszając objętość próbki — powodujemy spadek liczby istniejących w próbce neutronów, a równocześnie tendencję zaniku lub wręcz zanik reakcji łańcuchowej.

Zestawiając ostatnie wnioski powiemy, że rozmiary krytyczne próbek z czystego ^{235}U są niewielkie, próbek z czystego ^{238}U i uranu naturalnego natomiast — bardzo duże. Zmniejszenie rozmiarów krytycznych próbek z uranu naturalnego można osiągnąć wzbogacając je sztucznie w ^{235}U. Można wtedy dojść do rozmiarów krytycznych tego rzędu, co np. w owocu grejpfrutu.

Wyzwalanie olbrzymich ilości energii w reakcjach rozszczepienia znalazło zastosowanie: 1) w bombach atomowych — przy lawinowym niekontrolowanym przebiegu reakcji, 2) w reaktorach jądrowych — urządzeniach umożliwiających sterowane wyzwalanie energii. Budowie i działaniu reaktora poświęcony jest paragraf następny.

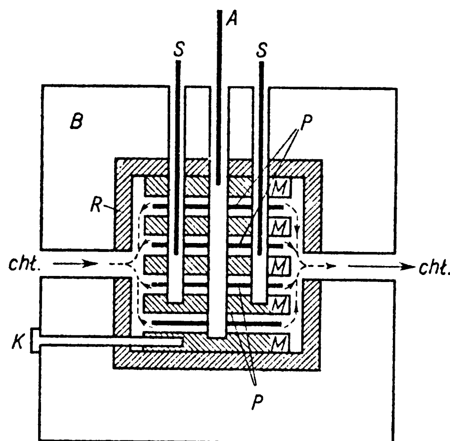
31.9. Reaktory jądrowe

Wiemy już, że użycie czystego ^{235}U, jako materiału rozszczepialnego, zapewniłoby lawinowy (wybuchowy) przebieg reakcji rozszczepienia. Taki przebieg jest oczywiście niepożądany, jeśli mamy na względzie wykorzystanie energii jądrowej dla celów pokojowych. W reaktorach jądrowych stosuje się jako tzw. paliwo jądrowe możliwie czysty uran naturalny, względnie uran sztucznie wzbogacony w izotop ^{235}U (do kilku lub kilkunastu procent).

Procesem podstawowym w reaktorze jest rozszczepienie ^{235}U, zachodzące najwydajniej w przypadku bombardowania neutronami termicznymi. A zatem w celu osiągnięcia wydajnej pracy reaktora należy odpowiednio spowolnić neutrony wyzwalane podczas rozszczepienia. Neutrony te jednak w swoim przebiegu przez masę paliwa ją-

drowego o dużej przewadze ^{238}U, zderzając się z ciężkimi jądrami ^{238}U, stopniowo nieznacznie zmniejszają swą prędkość, tak że jej wartość spada do wspomnianego już poprzednio rezonansowego przedziału prędkości. Przy takich prędkościach grozi neutronom pochłonięcie ich przez jądra ^{238}U, bez wywołania reakcji rozszczepienia. Te pochłonięte neutrony należy już „spisać na straty" w ogólnym bilansie reakcji łańcuchowej. Stąd widać, że w celu uzyskania wydajnej pracy reaktora należy neutrony szybko spowolnić, poniżej zakresu rezonansowych prędkości, aby zmniejszyć prawdopodobieństwo pasożytniczego wychwytu neutronu przez ^{238}U.

Pomysł szybkiego spowolnienia neutronów pochodzi od Fermiego i polega na wykorzystaniu zderzeń sprężystych neutronów z jądrami o zbliżonej masie (a więc np. z jądrami wodoru lekkiego lub ciężkiego, węgla, berylu). Jądra te występują we wprowadzonym w odpowiedni sposób do reaktora *spowalniaczu*, czyli *moderatorze prędkości*. Moderatorem prędkości może być woda lekka lub ciężka, grafit, niektóre związki berylu itp. Dla uzyskania odpowiednio efektywnego spowolnienia neutronów stosuje się paliwo jądrowe nie w postaci większej pojedynczej bryły, lecz w postaci oddzielnych tzw. prętów paliwowych. Pręty paliwowe mogą być po prostu zanurzone w ciekłym moderatorze lub umieszczone w kanałach wywierconych w moderatorze stałym.



Rys. 31.15

Poznaliśmy w ten sposób dwa podstawowe elementy występujące w każdym reaktorze uranowym, a mianowicie pręty paliwowe P i moderator prędkości M (rys. 31.15), umieszczone w środkowej części reaktora, zwanej rdzeniem R. Do regulowania szybkości przebiegu reakcji w reaktorze służą tzw. pręty sterujące S. Sporządzone są one z materiału zawierającego jądra atomowe wyróżniające się silną absorpcją neutronów, np. ze stali borowej lub kadmu. Głębsze ich zanurzenie w rdzeniu zmniejszając liczbę neutronów zwalnia przebieg reakcji.

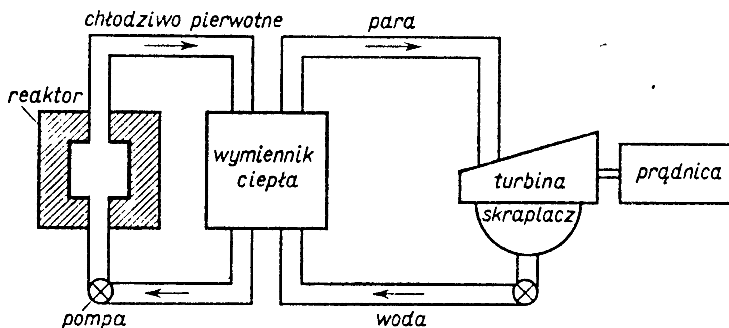
Prócz prętów paliwowych i sterujących każdy reaktor ma specjalne pręty awaryjne A, umożliwiające natychmiastowe przerwanie pracy reaktora. W razie awarii reaktora

pręty te automatycznie zostają zrzucone na dno rdzenia. Pochłaniają one tak intensywnie neutrony, że praktycznie reakcja przerywa się natychmiast.

Wiemy już, jak reakcja może się zakończyć, nie mówiliśmy natomiast nic o jej zainicjowaniu. Okazuje się, że zapoczątkowanie reakcji nie wymaga żadnych zabiegów: wystarczają neutrony swobodne, wszędzie istniejące, pochodzenia kosmicznego i neutrony wytwarzane podczas samorzutnego rozszczepienia uranu.

Interesujące mogą być dane dotyczące rozmiarów rdzenia reaktora. Mogą one być bardzo różne, zależnie od rodzaju paliwa i rodzaju moderatora. Bardzo duże rozmiary mają reaktory, których działanie opiera się na naturalnym uranie — jako paliwie i graficie — jako moderatorze. Na zmniejszenie rozmiarów wpływa: a) zwiększenie stężenia ^{235}U i b) zastąpienie grafitu np. wodą. Zawsze jednak rozmiary są bardzo duże, gdyż trzeba pamiętać o tym, że każdy reaktor musi mieć odpowiednie osłony ochronne i zabezpieczenie dodatkowe w postaci np. muru betonowego *B* o kilkumetrowej grubości. Jest to konieczne, ze względu na niebezpieczeństwo napromieniowania bardzo szkodliwymi promieniami γ i promieniami neutronowymi.

Podczas pracy reaktora wydzielają się w nim bardzo duże — wiemy już w jaki sposób regulowane — ilości energii. Ma ona postać energii kinetycznej produktów rozszczepienia, a z biegiem czasu ulega przemianie w energię cieplną. Ciepło wytwarzane w reaktorze jest odbierane przez tzw. chłodziwo (*chl* — na rys. 31.15) — substancję przepuszczaną przez reaktor. W zależności od rodzaju tej substancji mówi się o reaktorach z chłodzeniem wodnym, gazowym. Stosowane są także jako substancje chłodzące mieszaniny stopionych metali, takich jak sód, potas, mające dużą zaletę dobrej przewodności ciepła.



Rys. 31.16

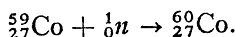
Energia cieplna wytworzona w reaktorze znalazła różne zastosowania. Ogrzane chłodziwo, np. wodne, może być wykorzystane do ogrzewania pomieszczeń. Na pierwsze miejsce jednak pod względem znaczenia wysuwają się zastosowania w energetyce, zilustrowane schematycznie na rys. 31.16. Chłodziwo pierwotne (np. stopiony metal) krąży w obiegu pierwotnym oddając w wymienniku ciepła swą energię wodzie. Woda zamieniona na parę krążąc w obiegu wtórnym służy do napędu turbogeneratorów elektrowni jądrowych lub silników, np. statków o napędzie jądrowym. Należy też wspom-

nieć o dużych walorach reaktorów z wrzącą wodą, w których woda zostaje doprowadzona do wrzenia bezpośrednio w reaktorze, a nie w wymienniku ciepła.

Oczywiście tak jak w każdym innym przypadku, przemiana energii cieplnej na energię mechaniczną w turbinach lub elektryczną w prądnicach odbywa się z dużymi stratami energii. Wydajność przemiany szacuje się na około 25%.

Znaczenie reaktorów jądrowych nie ogranicza się do wyżej wymienionych ich zastosowań. Reaktory odgrywają podstawową rolę w badaniach z dziedziny fizyki jądrowej. Wyniki badań są niejednokrotnie tego rodzaju, że pozwalają na rozszerzenie zastosowań na inne dziedziny, np. medycynę, technikę itp. Rozpatrzymy to nieco dokładniej.

Reakcje zachodzące w reaktorach atomowych są bardzo silnym źródłem neutronów, które, jak już wiemy, są najskuteczniejszymi pociskami wywołującymi przemiany jądrowe. W niektórych reaktorach ograniczony strumień neutronów jest wyprowadzany na zewnątrz przez odpowiednie „okienko" w ścianie reaktora i tam wykorzystywany do badań odpowiednich próbek. W innych — badaną próbkę wprowadza się do wnętrza reaktora przez odpowiedni kanał K i tam poddaje się silnemu napromieniowaniu neutronowemu. W wyniku napromieniowania powstają sztuczne izotopy promieniotwórcze, tzw. *radioizotopy*. Właśnie te — stosunkowo tanie — sztuczne izotopy promieniotwórcze znajdują zastosowanie np. w medycynie m. in. w walce z nowotworami (kobaltoterapia), w rolnictwie (np. otrzymywanie nowych odmian roślin uprawnych po naświetleniu nasion), w technice (np. modyfikowanie przebiegu procesu polimeryzacji, wulkanizacji różnych tworzyw, nie tylko kauczuku itp.). Szeroko znane jest zastosowanie sztucznego źródła promieniotwórczego, zwanego *bombą kobaltową*, otrzymywanego w reaktorze podczas działania neutronów na niepromieniotwórczy kobalt ^{59}Co:

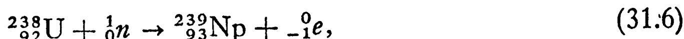


$$^{59}_{27}\mathrm{Co} + {}^{1}_{0}n \rightarrow {}^{60}_{27}\mathrm{Co}.$$

Promieniotwórczy kobalt ^{60}Co podczas rozpadu emituje promienie β, a równocześnie wysyła promienie γ — promienie o silnym działaniu jonizującym, niebezpieczne dla organizmu ludzkiego. Fakt ten narzuca oczywiście obowiązek stosowania odpowiednich środków ostrożności w czasie transportu bomby, jej przechowywania i stosowania do różnych celów.

31.10. Reaktory plutonowe

Z podanego wyżej opisu działania reaktora uranowego wynika, że w tym typie reaktora decydującą rolę odgrywa ^{235}U. Okazało się jednak, że i ^{238}U również może być bardzo przydatny dla celów wyzwalania energii jądrowej. Można go bowiem wykorzystać do otrzymywania dwóch pierwiastków transuranowych neptunu i plutonu:



$$^{238}_{92}\mathrm{U} + {}^{1}_{0}n \rightarrow {}^{239}_{93}\mathrm{Np} + {}^{0}_{-1}e, \tag{31.6}$$



$$^{239}_{93}\mathrm{Np} \rightarrow {}^{239}_{94}\mathrm{Pu} + {}^{0}_{-1}e. \tag{31.7}$$

Oba transurany są promieniotwórcze, lecz ich okresy połowicznego zaniku różnią się bardzo znacznie ($T_{Np} = 2{,}33$ dnia, $T_{Pu} = 24400$ lat). Pluton jest pierwiastkiem podle-

gającym reakcji rozszczepienia, może więc służyć jako materiał paliwowy w reaktorze. Ponieważ pluton ma właściwości chemiczne różne od uranu (nie jest izotopem uranu), można go więc wydzielić stosując szereg skomplikowanych procesów fizykochemicznych. W ten sposób w czasie drugiej wojny światowej wyprodukowano duże ilości tego pierwiastka zużywając go częściowo do produkcji drugiej bomby atomowej.

Powróćmy jednak do reaktora uranowo-plutonowego. W reaktorze tym stosuje się paliwo w postaci rdzenia z ^{235}U umieszczonego w otoczce z ^{238}U. Reakcja rozszczepienia zapoczątkowana w ^{235}U wyzwala szybkie neutrony, które — wobec braku moderatora prędkości w tym reaktorze — zgodnie z reakcjami (31.6) i (31.7) prowadzą do produkcji plutonu. Pluton $^{239}_{94}Pu$ ulega rozszczepieniu zarówno pod wpływem powolnych, jak i prędkich neutronów dostarczając nowych neutronów, które z kolei prowadzą znów do produkcji plutonu itd. Nowy typ reaktora ma zatem tę ważną zaletę, że sam w czasie pracy produkuje paliwo jądrowe. Ten typ reaktora zaliczamy do kategorii *reaktorów*: a) *rozmnażających* paliwo, b) *prędkich*, gdyż proces rozszczepienia jest w nim wywołany przez neutrony o dużej prędkości.

Rozdział ten zakończymy krótkimi uwagami na temat cząstek elementarnych. W dotychczasowych rozważaniach braliśmy pod uwagę protony, neutrony, elektrony, neutrina i antyneutrina, fotony. Nie wyczerpuje to jednak listy cząstek elementarnych. Badania promieni kosmicznych pozwoliły na wykrycie 2 odmian tzw. *mezonów* μ (*mionów*) o ładunku elementarnym $+e$ i $-e$ i masie 207 m_e, trzech odmian, tzw. *mezonów* π (*pionów*) o ładunku $+e$, $-e$ i zero i o masie bliskiej 270 m_e. Następnie odkryto tzw. *mezony K* (*kaony*) o masie około 970 m_e i tzw. *hiperony* — cząstki cięższe ponad 2000 razy od elektronów. Wszystkie wymienione cząstki są nietrwałe.

Wiele wiadomości o nowych cząstkach elementarnych dostarczają badania, w których wykorzystuje się działanie protonów lub cząstek α, rozpędzanych w nowoczesnych akceleratorach. Uderzenie takiej cząstki o energii rzędu gigaelektronowoltów nie przekształca już uderzonego jądra w inne, lecz wybija z niego szereg cząstek elementarnych. Badanie ich właściwości i poszukiwanie nowej koncepcji struktury materii — oto cel, jaki sobie wytyczają liczni współcześni fizycy.

Pytania i zadania

1. Jakie doświadczenia przemawiały na korzyść modelu atomu Rutherforda? Omówić je dokładnie
2. Jaka jest ogólna charakterystyka jąder atomowych z punktu widzenia: a) rozmiarów, b) struktury.
3. Podać definicję nukleonów.
4. Jak należy odczytać symbol ^{A}X?
5. Na czym polega izotopia?
6. Co decyduje o trwałości jąder atomowych? Co nazywamy energią wiązania? Jak się zmienia energia wiązania w zależności od liczby masowej A?
7. Jakie siły występują w jądrze atomowym? Jaka jest zasadnicza różnica między nimi?
8. Na czym polega promieniotwórczość naturalna? W jaki sposób można rozdzielić składniki promieniowania α, β i γ? Podać zasadnicze różnice wymienionych promieni. Czy wszystkie trzy rodzaje promieni emitowane są z jądra tego samego określonego pierwiastka?
9. Jakie są szeregi (rodziny) promieniotwórcze? Jak można je ogólnie scharakteryzować?

10. Jakie prawa rządzą kinetyką przemian promieniotwórczych? Podać definicje: a) stałej rozpadu, b) okresu połowicznego zaniku?

11. Dlaczego w przypadku promieniotwórczości naturalnej mówimy o widmie prędkości cząstek β, natomiast nie mówimy o widmie prędkości cząstek α? Jak się tłumaczy istnienie neutrina?

12. Jakie są podstawowe metody detekcji cząstek naładowanych? Omówić zasadę budowy i działania najważniejszych detektorów.

13. Jakie czynniki mogą wywołać reakcje jądrowe?

14. Jakie urządzenia objęte są ogólną nazwą akceleratorów? Jakie są typy akceleratorów? Podać zasady budowy i działania akceleratorów różnych typów.

15. Na czym polegają reakcje rozszczepienia jąder atomowych? Dlaczego reakcji rozszczepienia towarzyszy wyzwalanie dużych ilości energii?

16. Wyjaśnić koncepcję modelu kroplowego jądra atomowego. Podać przykłady zjawisk, które można wytłumaczyć na podstawie tego modelu.

17. Na czym polega reakcja łańcuchowa i lawinowa?

18. Czy słuszne jest twierdzenie, że w 50-tej generacji idealnej reakcji rozszczepienia, w której w każdym etapie wyzwalane są dwa neutrony, powstanie około 10^{15} neutronów?

19. Dlaczego rzeczywiste reakcje rozszczepienia mają przebieg bardzo różny od przebiegu idealnego? Jakie zjawiska mogą następować po wyzwoleniu neutronu podczas rozszczepienia ^{238}U i ^{235}U? Czy wszystkie efekty są równie prawdopodobne w obu wymienionych izotopach uranu?

20. Omówić budowę i działanie reaktorów jądrowych podkreślając rolę i znaczenie każdego elementu.

21. Omówić podstawowe zastosowania reakcji łańcuchowych.

22. Znaleźć stałą rozpadu radu, jeśli okres połowicznego zaniku wynosi 1550 lat.

Odp. $1{,}42 \cdot 10^{-11}$ s^{-1}.

23. Licząc scyntylacje towarzyszące rozpadowi 1 g ^{238}U otrzymano, że na sekundę przypada ich $1{,}15 \cdot 10^4$. Obliczyć na tej podstawie stałą rozpadu ^{238}U.

Odp. $45 \cdot 10^{-19}$ s^{-1}.

24. Obliczyć energię unoszoną przez cząstki α powstające w wyniku rozpadu 1 g radu w ciągu godziny, jeśli prędkość cząstek równa się $1{,}51 \cdot 10^7$ m/s, a okres połowicznego zaniku radu równa się 1550 lat. Dalszych produktów rozpadu nie brać pod uwagę.

Odp. 99,5 J.

25. Jaką objętość w warunkach normalnych będzie zajmował hel wydzielony w wyniku rozpadu 1 g radu w ciągu 1 roku, jeśli wiadomo, że liczba rozpadów 1 g radu w ciągu sekundy wynosi $3{,}7 \cdot 10^{10}$. Dalszych produktów rozpadu nie brać pod uwagę.

Odp. ok. 43 mm^3.

26. Na rysunku 31.3 podany jest schemat kolejnych reakcji w szeregu uranowo-radowym. Obliczyć liczbę cząstek α emitowanych w ciągu jednej sekundy przez mieszaninę 1 mg ^{226}Ra i produktów jego rozpadu do RaD włącznie, z którymi pozostaje w równowadze promieniotwórczej. Okres połowicznego zaniku wynosi $T_{Ra} = 1590$ lat. (U w a g a: Warunki równowagi promieniotwórczej charakteryzują się tym, że nie zmieniają się ilości poszczególnych produktów rozpadu.)

Odp. $1{,}48 \cdot 10^8$ s^{-1}.

27. Wiedząc, że okres połowicznego zaniku pewnego pierwiastka $T = 20$ dni obliczyć: 1) po jakim czasie rozpadnie się 3/4 początkowej liczby atomów, 2) po jakim czasie pozostanie 1/8 początkowej liczby atomów, 3) ile wynosi stała rozpadu.

Odp. 1) 40 dni, 2) 60 dni, 3) około $4 \cdot 10^{-7}$ s^{-1}.

28. W blendzie uranowej znaleziono izotop ^{206}Pb, co wskazuje na promieniotwórcze pochodzenie ołowiu. Jaki jest wiek blendy, jeżeli na każdy gram ^{238}U przypada 0,80 g ^{206}Pb? Okres połowicznego zaniku uranu $T = 4{,}5 \cdot 10^9$ lat.

Odp. $4{,}25 \cdot 10^9$ lat.

ROZDZIAŁ 32

Fizyka ciała stałego

32.1. Uwagi wstępne

Ostatni rozdział niniejszego podręcznika poświęcamy wybranym zagadnieniom z tzw. fizyki ciała stałego, działu fizyki współczesnej rozwijającego się bardzo szybko w ciągu ostatnich kilkudziesięciu lat.

O strukturze ciał stałych mówiliśmy już sporo w rozdziale 16, rozważając charakterystyczną dla tych ciał budowę krystaliczną z punktu widzenia teorii kinetyczno-molekularnej. Podkreślaliśmy regularne i okresowo powtarzające się rozmieszczenie atomów, jonów czy cząsteczek w obrębie badanego ciała, odpowiedzialne za istnienie takiego lub innego typu sieci krystalicznej.

O możliwościach badania struktury krystalicznej ciał mówiliśmy w rozdziale 31 przy okazji omawiania różnych metod analizy rentgenowskiej. Wiemy już, że wiele informacji można uzyskać na podstawie rozmaitego typu rentgenogramów.

Jednak wiele zagadnień dotyczących struktury wewnętrznej, a przede wszystkim związku struktury z właściwościami fizykochemicznymi ciał wymagało bliższego zbadania. Między innymi wyjaśnienia wymagała często obserwowana niezgodność wniosków wyciąganych z teorii kryształów doskonałych z faktami stwierdzonymi doświadczalnie w przypadku kryształów rzeczywistych. Przyczyn tej niezgodności należało szukać we wpływie różnego rodzaju defektów krystalicznych (domieszek, dyslokacji itp.) na właściwości ciał stałych. Między innymi i o tych sprawach będziemy mówili w niniejszym rozdziale.

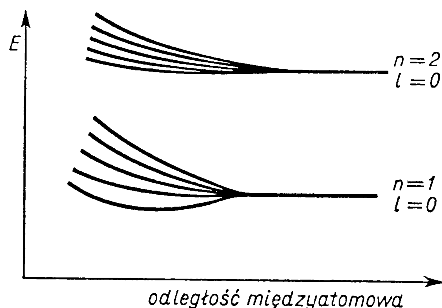
Rozważania nasze, dotyczące fizyki ciała stałego, rozpoczniemy od stwierdzenia, że w przypadku ciała stałego mamy do czynienia z materią w postaci skondensowanej, to znaczy z takim stanem materii, w którym wzajemne odległości między cząsteczkami danego ciała są małe w porównaniu z wymiarami poszczególnych cząsteczek. Jak się przekonamy, w tym przypadku konieczne będzie wprowadzenie zasadniczych zmian do przedstawionego w rozdziale 30 obrazu struktury elektronowej izolowanych atomów.

Najistotniejsze wyniki rozważań dotyczących izolowanego atomu można ująć w ten sposób: 1) elektronom w atomie przypisujemy pewne dyskretne, dozwolone stany

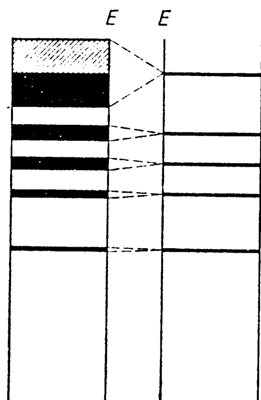
energetyczne, 2) przyjmujemy jako obowiązującą zasadę Pauliego, określającą maksymalną liczbę elektronów w poszczególnych powłokach i podpowłokach atomu.

Przechodząc do właściwości ciała stałego weźmy początkowo pod uwagę zamiast izolowanego atomu układ złożony ze skończonej, niewielkiej liczby atomów. Niech to będzie np. układ złożony z 5 atomów wodoru umieszczonych w stałych odległościach od siebie. Elektron każdego z tych atomów podlega działaniu sił pochodzących od jego bliższego i dalszego otoczenia.

Przedstawiając na wykresie zależność możliwych stanów energetycznych każdego z elektronów od odległości międzyatomowej otrzymalibyśmy wykres taki jak na rys. 32.1. W przypadku dużych odległości międzyatomowych atomy są praktycznie od siebie niezależne. Otrzymujemy pojedyncze poziomy energetyczne zmieniające się wraz ze zmianą liczb kwantowych. Gdy jednak odległości między atomami maleją, to zamiast pojedynczych poziomów energetycznych otrzymujemy poziomy rozszczepione i to tym bardziej, im odległości międzyatomowe są mniejsze.



Rys. 32.1



Rys. 32.2

Rysunek 32.1 charakteryzuje w sposób poglądowy układ złożony z pięciu atomów wodoru. Sytuacja jest oczywiście znacznie bardziej złożona, gdy mamy do czynienia z gęstym zbiorem atomów w ciele stałym. Zakłócenia poziomów energetycznych (najsilniejsze dla elektronów w najbardziej zewnętrznych powłokach, a więc dla elektronów walencyjnych), wynikają z oddziaływania wielkiej liczby różnie rozmieszczonych pozostałych jąder atomowych i elektronów w ciele stałym. Zamiast pojedynczego poziomu energetycznego pojawia się układ bardzo bliskich siebie poziomów energetycznych, quasi-pasmo złożone z bardzo wielu bliskich siebie dozwolonych stanów energetycznych.

Jeden dyskretny poziom energetyczny charakterystyczny dla izolowanego atomu (rys. 32.2 — prawa strona) zostaje zastąpiony przez continuum wartości energii w pewnym określonym przedziale (pasmie) (rys. 32.2 — lewa strona). Takie *pasmo dozwolone* jest oddzielone od następnego pasma dozwolonego *pasmem zabronionym* o mniejszej lub większej szerokości w skali energii. W obszarze jednego pasma elektrony mogą łatwo

przechodzić z jednego poziomu energetycznego na drugi, gdyż różnice są niewielkie. Jest tylko jeden ważny warunek: nie może być przekroczony zakaz Pauliego. Jeśli w danym pasmie wszystkie stany energetyczne są obsadzone elektronami zgodnie z zasadą Pauliego, to ruchu elektronów nie będzie. O takim pasmie mówimy, że jest *pasmem zapełnionym*.

Jeżeli jednak dane pasmo jest tylko częściowo obsadzone elektronami, to możliwe są przejścia elektronów na poziomy energetyczne wyższe i takie pasmo nosi nazwę *pasma przewodnictwa*. Te pojęcia — podstawowe dla tak zwanego modelu pasmowego ciała stałego — wykorzystamy w § 32.8 do charakterystyki metali, izolatorów (dielektryków) i półprzewodników. Zanim jednak przejdziemy do tej charakterystyki, ustalimy podstawowe dane dotyczące struktury elektronowej typowych pierwiastków metalicznych i zestawimy wnioski, jakie można wyciągnąć z modelu struktury metalu opartego na koncepcji elektronów swobodnych.

32.2. Struktura elektronowa atomów typowych pierwiastków metalicznych

Przedstawienie wyczerpujących danych dotyczących istniejących w przyrodzie metali wykracza oczywiście poza ramy niniejszego podręcznika. Mówiąc na przykład o strukturze elektronowej atomów metali ograniczymy się do krótkiej charakterystyki atomów metali jednowartościowych, należących do grupy metali alkalicznych (lit, sód, potas, rubid, cez) oraz do grupy metali szlachetnych (miedź, srebro, złoto).

Tabela 30.1 dostarcza danych dotyczących struktury elektronowej Li ($Z = 3$), Na ($Z = 11$) i K ($Z = 19$). Korzystając z obszerniejszej tablicy, obejmującej dalsze pierwiastki układu okresowego, uzyskalibyśmy dane dotyczące struktury elektronowej Rb ($Z = 37$) i Cs ($Z = 55$). Stwierdzilibyśmy, że we wszystkich przypadkach mamy do czynienia z jednym elektronem walencyjnym należącym do podpowłoki o liczbie kwantowej $l = 0$. Podpowłoka ta zaczyna się rozbudowywać po wypełnieniu podpowłoki o liczbie kwantowej $l = 1$ należącej do poprzedniej powłoki. Stwierdzenie to nie dotyczy oczywiście litu, w którego atomie nie ma w ogóle podpowłoki o liczbie kwantowej $l = 1$. Wspomniane elektrony walencyjne są to elektrony: $2s$ w Li, $3s$ w Na, $4s$ w K, $5s$ w Rb i $6s$ w Cs.

Analogiczny przegląd struktury elektronowej metali szlachetnych prowadzi do następujących wniosków. Jedyny elektron walencyjny atomów tych metali znajduje się we wszystkich przypadkach w podpowłoce o liczbie kwantowej $l = 0$ (a więc analogicznie do metali alkalicznych), z tym jednak, że rozbudowa tej podpowłoki rozpoczyna się po zapełnieniu podpowłoki o $l = 2$ w powłoce poprzedniej.

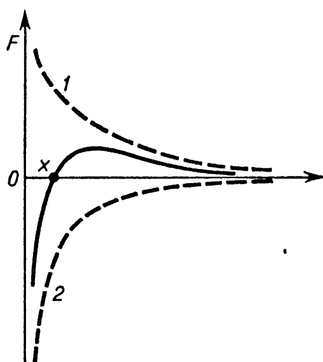
32.3. Wiązania metaliczne

Już w rozdziale 16 wprowadziliśmy podstawowe pojęcia z krystalografii, które obecnie wykorzystamy przy omawianiu struktury krystalicznej metali.

Metale krystalizują najczęściej w jednym z trzech układów: regularnym scentrowanym przestrzennie (np. Li, Na, K, Cs), regularnym scentrowanym powierzchniowo

(np. Cu, Ag, Au, Pt, Ni, Al) i heksagonalnym o najgęstszym upakowaniu (np. Be, Mg, Zn, Cd).

Rozważając zagadnienie sił utrzymujących atomy w krysztale w równowadze trzeba brać pod uwagę zależność sił międzyatomowych od odległości. Należy tu rozróżnić siłę przyciągającą, malejącą szybko ze wzrostem odległości (krzywa kreskowana *1* na rys. 32.3) oraz siłę odpychającą (krzywa kreskowana *2* na tym samym rysunku) gwałtownie rosnącą po przekroczeniu pewnej granicznej odległości. Ten gwałtowny wzrost siły odpychania wiąże się z zakazem Pauliego. A mianowicie, gdy odległość 2 atomów zmniejszy się tak dalece, że orbity elektronowe obu atomów mogą zachodzić na siebie, pojawia się możliwość wystąpienia w danym atomie 2 elektronów o wszystkich jednakowych liczbach kwantowych. Szybko w tych warunkach rosnące siły odpychania tworzą fizyczne podstawy zakazu Pauliego. Równocześnie rośnie siła przyciągania elektronów walencyjnych przez pobliskie rdzenie jądrowe.



Rys. 32.3

Dla pewnej odległości międzyatomowej x siły przyciągania i odpychania wzajemnie się równoważą (krzywa ciągła przedstawiająca siłę wypadkową przechodzi przez zero). Właśnie takie warunki stwarzają możliwość formowania się struktury krystalicznej.

32.4. Charakterystyka właściwości metali oparta na koncepcji elektronów swobodnych

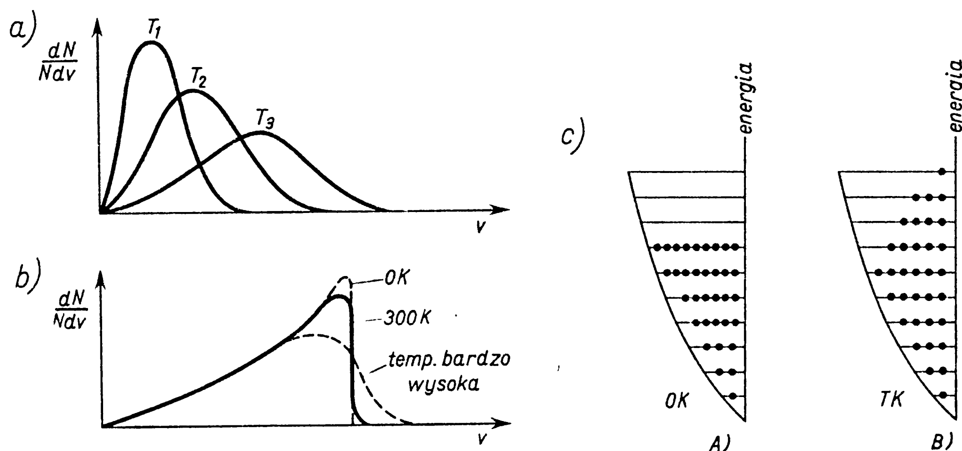
Pojęciem elektronów swobodnych w metalu posługiwaliśmy się już wielokrotnie, gdyż jest ono bardzo przydatne do wyjaśniania licznych właściwości metali.

Nazwą elektrony swobodne obejmujemy elektrony walencyjne, które mogą poruszać się swobodnie w całej objętości metalu. Są one odpowiedzialne za zjawisko przewodnictwa elektrycznego i dlatego często nazywane są *elektronami przewodnictwa*, w odróżnieniu od elektronów należących do zapełnionych powłok elektronowych. W modelu struktury metalu opartym na swobodnych elektronach zaniedbuje się oddziaływania między elektronami przewodnictwa i ich jonowo-elektronowym otoczeniem,

czyli resztą sieci krystalicznej metalu, co oczywiście jest wielkim uproszczeniem warunków rzeczywistych.

Można tu przypomnieć kilka osiągnięć, a następnie niepowodzeń, jakie towarzyszyły stosowaniu wspomnianego modelu. W § 20.10 podaliśmy klasyczne wyprowadzenie prawa Ohma w ujęciu teorii elektronowej, mówiliśmy też o zależności Wiedemanna-Franza wiążącej prawidłowo przewodnictwo elektryczne metalu z przewodnictwem cieplnym w temperaturach bliskich pokojowej. W ujęciu klasycznym swobodne elektrony traktowane były jako gaz elektronowy, do którego można stosować podstawowe wzory teorii kinetycznej gazów. Przyjmowano, między innymi, że rozkład prędkości w gazie elektronowym podlega prawu rozkładu prędkości Maxwella (tzw. *statystyka Maxwella–Boltzmanna*).

Wspomniany model zawodzi między innymi przy ustalaniu ciepła molowego i właściwego metali. Według teorii kinetycznej, licznie w metalu występujące swobodne elektrony dzięki swemu ruchowi postępowemu (analogicznemu do ruchu postępowego jednoatomowych gazów — 3 stopnie swobody) powinny mieć duży wpływ na wartość ciepła molowego, a tego nie stwierdza się doświadczalnie. Trudności wypływają również wtedy, gdy uwzględniając istnienie momentu magnetycznego każdego swobodnego elektronu poszukuje się w metalach umieszczonych w słabym polu magnetycznym dość silnego efektu paramagnetycznego (por. § 21.8). Efekt taki w przewodnikach metalicznych nie występuje.



Rys. 32.4

Niektóre ze wspomnianych trudności udało się usunąć zastępując klasyczny model metalu ze swobodnymi elektronami modelem kwantowym. Zmiana polega na tym, że odrzuca się założenie ciągłego rozkładu prędkości i ciągłego rozkładu energii związane ze statystyką Maxwella–Boltzmanna, a wprowadza się *kwantową statystykę Fermiego–Diraca*, uwzględniającą zakaz Pauliego. Aby dokładniej zrozumieć różnice, skorzystamy z graficznego przedstawienia stosunków odpowiadających obu statystykom. Rysunek 32.4a przedstawia przebieg krzywych rozkładu prędkości Maxwella dla trzech

temperatur: $T_1 < T_2 < T_3$. Widać wyraźnie stopniowe spłaszczanie się krzywych i wydłużanie się ich w kierunku większych prędkości. W ujęciu tej statystyki w temperaturze zera bezwzględnego średnia prędkość cząstek równa się zeru.

Rysunek 32.4b przedstawia odpowiedni rozkład Fermiego–Diraca również dla trzech różnych temperatur, a rys. 32.4c przedstawia czysto poglądowo obsadzenie dozwolonych skwantowanych poziomów energetycznych w temperaturze 0K (rys. A) i wyższej (rys. B).

W temperaturze 0K elektrony mają tendencję do zajęcia możliwie najniższych poziomów energetycznych, ale wobec założenia, że obowiązuje zakaz Pauliego siłą rzeczy część elektronów zajmuje poziomy energetycznie wyższe. Stąd wynika ważny wniosek: w temperaturze 0K energia kinetyczna elektronów nie tylko nie równa się zeru, ale nawet ich średnia energia kinetyczna znacznie różni się od zera. Wraz ze wzrostem temperatury część elektronów z najwyższych obsadzonych w temperaturze 0K poziomów przechodzi do wyższych skwantowanych stanów energetycznych. W temperaturach niezbyt wysokich przebieg krzywych rozkładu Fermiego–Diraca mało się różni od przebiegu tychże krzywych dla temperatury 0K.

Kilku słów uzupełnienia wymaga kilkakrotnie już wspomniane założenie, że do elektronów przewodnictwa w metalu stosujemy zasadę Pauliego. Umiemy ją stosować w odniesieniu do izolowanych atomów pierwiastków; na jej podstawie wyjaśniliśmy budowę układu okresowego pierwiastków.

Obecnie mamy sytuację nową, zamiast zespołu elektronów zajmujących kolejne powłoki i podpowłoki atomu mamy w metalu elektrony swobodne, które też chcemy zgrupować w pewnych dozwolonych stanach dobierając odpowiednio wartości układu liczb kwantowych. Mechanika falowa pozwala na ustalenie nowych czterech liczb kwantowych określających stan swobodnego elektronu. Zastosowanie do nich zasady Pauliego prowadzi do wniosku, że poziomowi energetycznemu podstawowemu odpowiadają 2 poziomy dozwolone, które mogą być obsadzone przez 2 elektrony, pierwszemu poziomowi wzbudzonemu odpowiada 12 poziomów, a więc ich obsadę stanowi 12 elektronów. Drugi poziom wzbudzony ma maksymalną liczbę stanów do obsadzenia równą 24 itd. W temperaturze zera bezwzględnego elektrony stopniowo zapełniają kolejne poziomy licząc od podstawowego. Wreszcie po rozmieszczeniu wszystkich elektronów swobodnych na kolejnych poziomach osiągamy taką sytuację w temperaturze zera bezwzględnego, że pewien poziom energetyczny rozgranicza poziomy całkowicie zapełnione od całkowicie pustych. Ten rozgraniczający poziom nazywamy *poziomem Fermiego* w 0K, a energię mu odpowiadającą — energią Fermiego w 0K.

W tabeli 32.1 podane są wartości energii Fermiego dla kilku metali. Na przykład dla sodu energia Fermiego E_F równa się 3,12 eV, a średnia energia kinetyczna elektronów sodu (równa, jak można wykazać $\frac{3}{5}E_F$) wynosi 1,87 eV. Dotyczy to temperatury zera bezwzględnego. Dla porównania warto może podać, że do osiągnięcia takiej średniej energii kinetycznej cząstek, gdy obowiązuje statystyka Maxwella, potrzebna jest temperatura około 14 000K. W ten sposób raz jeszcze podkreśliliśmy zasadniczą różnicę w ujęciu klasycznym i kwantowym. Według ujęcia klasycznego w temperaturze zera bezwzględnego wszystkie elektrony swobodne powinny się znajdować w stanie

energetycznym zerowym, według ujęcia kwantowego zapełniają one wszystkie stany energetyczne dozwolone, aż do wartości około 5 eV (tab. 32.1) w stosunku do stanu podstawowego.

Tabela 32.1

Energie Fermiego (w 0K) oraz prace wyjścia czystych metali

Pierwiastek	E_F eV	W_w eV
lit	4,7	2,2
miedź	7,0	4,1
potas	2,1	1,8
sód	3,1	1,9
srebro	5,5	4,7
wolfram	5,8	4,5

Po wprowadzeniu modelu kwantowego, opartego na koncepcji elektronów swobodnych, powróćmy do trudności występujących przy wyjaśnianiu pewnych faktów na podstawie modelu klasycznego i przeanalizujmy np. raz jeszcze efekty związane z ogrzewaniem metalu z punktu widzenia obu modeli. W ujęciu klasycznym ogrzanie ciała od temperatury zera bezwzględnego do temperatury TK wywołuje wzrost energii kinetycznej każdej cząstki o wartość rzędu kT (dla temperatury pokojowej ok. 0,02 eV). W ujęciu kwantowym jest inaczej, jak już bowiem sygnalizowaliśmy, dozwolone stany energetyczne w temperaturze zera bezwzględnego są wypełnione do poziomu Fermiego, a więc do około 5 eV. Podniesienie na wyższy dozwolony poziom energetyczny przy dostarczaniu wskutek ogrzewania energii około 0,02 eV może dotyczyć tylko niewielkiego procentu ogólnej liczby elektronów swobodnych, a mianowicie tych elektronów, które zajmują stany energetyczne poniżej poziomu Fermiego w granicach rzędu kT. Jedynie one mogą być pobudzone termicznie, czyli tylko one mają wpływ na wartość ciepła właściwego (ciepła molowego) metalu. I tylko te nieliczne wysokoenergetyczne elektrony poruszają się swobodnie pod wpływem przyłożonego gradientu potencjału lub temperatury, a więc są odpowiedzialne za przewodnictwo cieplne i elektryczne w przewodnikach.

32.5. Rozkład potencjału wewnątrz i na powierzchni metalu

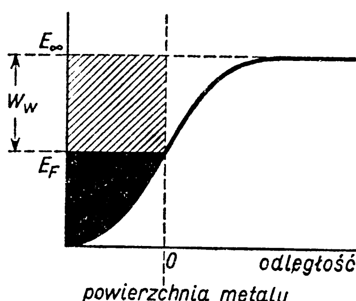
Rozwijając dalej koncepcję modelu metalu z elektronami swobodnymi możemy odpowiedzieć na pytanie, dlaczego elektrony swobodne poruszające się bez przeszkód wewnątrz metalu nie wychodzą na zewnątrz. Metal w ujęciu makroskopowym jest elektrycznie obojętny. Wyjście elektronu na zewnątrz powoduje dodatnie naładowanie metalu, co z kolei wywołuje powstanie pola elektrycznego. Innymi słowy, w pobliżu powierzchni metalu powstaje „bariera" potencjału (rys. 32.5) utrudniająca wyjście elektronu na zewnątrz. Oczywiście barierę tę mogą pokonać elektrony mające dosta-

tecznie duży zasób energii, uzyskany np. dzięki naświetlaniu odpowiednim rodzajem promieniowania (fotoemisja lub inaczej zjawisko fotoelektryczne, por. § 29.7) lub też dzięki odpowiedniemu ogrzaniu (termoemisja por. § 32.6). Dla każdego metalu istnieje charakterystyczna energia progowa, zwana inaczej *pracą wyjścia*, potrzebna do pokonania bariery potencjału.

O pracy wyjścia mówiliśmy już przy okazji omawiania zjawiska fotoelektrycznego. Z punktu widzenia modelu metalu z elektronami swobodnymi zdefiniujemy ją jako pracę potrzebną do przeniesienia elektronu z poziomu Fermiego o energii E_F do nieskończoności. Jeśli zatem pracę przeniesienia do nieskończoności elektronu z najniższego stanu energetycznego oznaczymy przez E_∞, to praca wyjścia W_w będzie równa

$$W_w = E_\infty - E_F.$$

Kilka wartości W_w podaje tabela 32.1. Do zagadnienia emisji elektronów z metalu wrócimy w § 32.6.



Rys. 32.5

Następne pytanie dotyczy rozkładu potencjału wewnątrz metalu. Ujmując sprawę makroskopowo wobec regularnej budowy metalu można uważać, że wnętrze metalu pozbawione jest pola elektrycznego. Dzięki takiemu założeniu elektrony walencyjne w metalu (bez przyłożenia pola zewnętrznego) odpowiadają cząsteczkom gazu doskonałego (bez oddziaływań międzycząsteczkowych). Ujmując sprawę mikroskopowo stwierdzimy niejednorodność potencjału wewnątrz metalu, gdyż każde jądro atomowe — aczkolwiek ekranowane (osłonięte) wewnętrznymi, związanymi elektronami — wywiera jednak działanie na elektrony walencyjne.

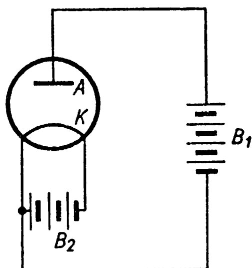
Uwzględnienie oddziaływania sieci krystalicznej na elektrony walencyjne prowadzi, jak już wiemy, do innego modelu struktury ciała stałego, a mianowicie do modelu pasmowego, pozwalającego rozróżnić grupę metali, izolatorów i półprzewodników. Omówimy go bardziej szczegółowo w § 32.8.

32.6. Emisja elektronów z metalu

Jak już mówiliśmy, zjawisko emisji elektronów z powierzchni metalu możliwe jest wtedy, gdy elektron ma zasób energii równy co najmniej pracy wyjścia. Podstawowymi sposobami nadawania elektronom żądanej energii są: ogrzanie i naświetlenie prowa-

dzące do termoemisji i fotoemisji. Prócz tych rodzajów emisji istnieje jeszcze emisja tak zwana polowa (lub zimna) i emisja tak zwana wtórna, które pokrótce scharakteryzujemy. *Emisja polowa* zachodzi wtedy, gdy w otoczeniu powierzchni metalu istnieje silne pole elektryczne o kierunku sprzyjającym wychodzeniu elektronów z powierzchni metalu. Efekt przyłożenia takiego pola sprowadza się do obniżenia bariery potencjału na granicy metalu. *Emisja wtórna* występuje wtedy, gdy powierzchnia metalu jest bombardowana elektronami lub jonami o dużej energii. Jeśli energia bombardujących elektronów jest duża (rzędu kilkuset eV), liczba emitowanych wtórnych elektronów jest większa od liczby elektronów padających. Zjawisko to zostało wykorzystane w różnego typu wzmacniaczach, między innymi w znanym nam już fotopowielaczu.

Zjawisko *fotoemisji* omawialiśmy szeroko w § 29.7. Obecnie zajmiemy się termoemisją.



Rys. 32.6

Zjawisko *termoemisji* rozpatruje się zwykle jako jedną z możliwości otrzymywania prądu elektrycznego w wysokiej próżni. Wyobraźmy sobie drucik metalowy K wprowadzony do bańki próżniowej, połączony ze specjalną baterią żarzenia B_2 (rys. 32.6). Przypuśćmy, że ma on potencjał ujemny (a więc stanowi katodę) w stosunku do drugiej elektrody A — anody połączonej z biegunem dodatnim baterii anodowej B_1. W takim urządzeniu anoda „zbiera" elektrony wysyłane przez rozżarzony drucik, podtrzymując przepływ ciągłego prądu termoemisji. Łatwo się domyślić, że natężenie tego prądu zależy od pracy wyjścia metalu katody, jak również — i to bardzo silnie — od temperatury katody. Omówione urządzenie stanowi znaną ogólnie dwuelektrodową lampę katodową (por. Uzupełnienie U.3). Wartość pracy wyjścia z danego metalu decyduje o tym, przy jakiej temperaturze metalu rozpoczyna się termoemisja. Różnice są bardzo duże, na przykład dla wolframu i platyny temperatury te są rzędu 2100–2300K, a dla wapnia, strontu i baru wynoszą około 800K.

Warto podkreślić pewną analogię między zjawiskiem termoemisji i zjawiskiem parowania. Załóżmy, że ogrzewanie drucika w próżni odbywa się bez włączenia pola elektrycznego między anodą i katodą. W otoczeniu żarzącego się drucika ustala się pewien stan równowagi. Tworzy się ładunek przestrzenny w postaci chmury elektronowej, z której elektrony wracają do katody i do której wchodzą coraz to nowe emitowane z drucika elektrony. Podobnie, w otoczeniu powierzchni swobodnej cieczy istnieje stan równowagi: z powierzchni cieczy ulatniają się cząsteczki o odpowiednio

dużej energii kinetycznej i odpowiednio skierowanej prędkości, a równocześnie cząsteczki pary z otoczenia powierzchni swobodnej cieczy w swoim bezładnym ruchu cieplnym wracają do obszaru cieczy. Im wyższa jest temperatura cieczy, tym gwałtowniej odbywa się proces parowania. Podobnie, im wyższa jest temperatura żarzenia drucika, tym więcej elektronów w jednostce czasu opuszcza jego powierzchnię.

Włączenie napięcia między anodą i katodą powoduje, jak już mówiliśmy, przepływ prądu termoemisji w przestrzeni między elektrodami. Ze wzrostem napięcia (przy danej temperaturze żarzenia) natężenie płynącego prądu rośnie do pewnej granicy. Maksymalną wartość prądu nazywamy *prądem nasycenia*. Z prądem nasycenia mamy do czynienia wtedy, gdy wszystkie elektrony emitowane w jednostce czasu z powierzchni katody są w tym samym czasie ściągane do anody. Można też mówić o gęstości prądu nasycenia rozumiejąc przez to ilość ładunków emitowanych w jednostce czasu przez jednostkę powierzchni żarzonej katody.

Gęstość prądu nasycenia dla określonego materiału katody jest zależna od temperatury. Wykorzystując kwantowy model metalu ze swobodnymi elektronami otrzymano następujący wzór na gęstość j prądu nasycenia w zjawisku termoemisji:

$$j = AT^2 e^{-W_w/kT},$$

gdzie A oznacza stałą $= \frac{4\pi mk^2 e}{h^3} = 120{,}4 \frac{A}{m^2 \cdot K^2}$, T — temperaturę w K, W_w — pracę wyjścia, k — stałą Boltzmanna $= 1{,}38 \cdot 10^{-23}$ J/K.

Teoretyczna wartość stałej A dość znacznie (2–8-krotnie) różni się od wartości eksperymentalnych otrzymanych dla różnych metali. Dla wyjaśnienia tych rozbieżności przyjmuje się, że pewien procent elektronów, mających w zasadzie odpowiednio dużą energię termiczną do wykonania pracy wyjścia, ulega jednak rozproszeniu na powierzchni metalu i wraca do jego wnętrza.

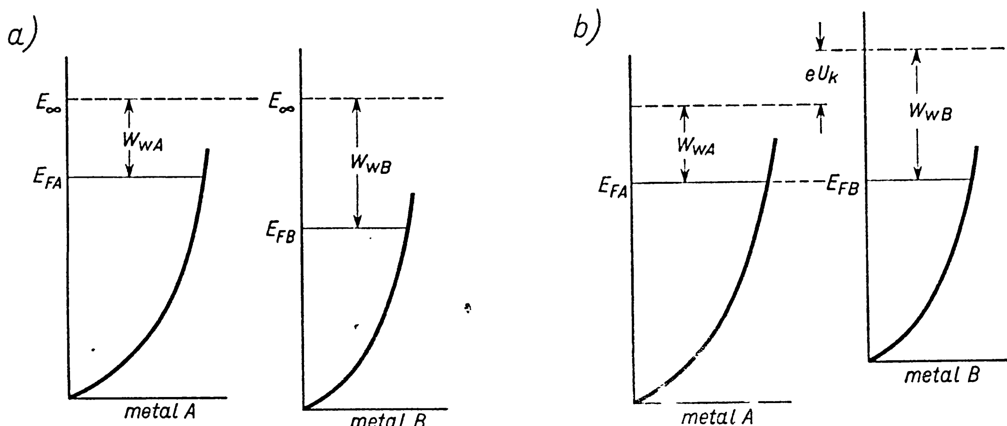
Wbrew oczekiwaniom czyste metale o najniższej pracy wyjścia, np. Cs, K, Na, Li nie znalazły szerokiego zastosowania jako emitery termoelektronów, a to z tego względu, że ich temperatury topnienia są stosunkowo niskie i w związku z tym przy silnym ogrzaniu szybko parują i ulegają zniszczeniu. Najszerzej stosowany w praktyce jest wolfram ($W_w = 4{,}52$ eV, temperatura topnienia 3643 K) zresztą nie w postaci czystej, lecz z domieszkami, które mają bardzo duży wpływ na termoemisję.

Zjawisko termoemisji znalazło szerokie zastosowanie przede wszystkim w różnego typu lampach katodowych (por. Uzupełnienia, U.3), oscylografach, w lampach rentgenowskich próżniowych itp.

32.7. Potencjały kontaktowe. Zjawisko termoelektryczne

Zjawisko potencjałów kontaktowych wynika z różnic wartości prac wyjścia i poziomów Fermiego dla różnych metali. Wyobraźmy sobie 2 metale A i B: metal A o wyższym poziomie energii Fermiego E_{FA} i mniejszej pracy wyjścia W_{wA}, jak to widać z rys. 32.7a. Po połączeniu obu metali elektrony z najwyższych dozwolonych poziomów energetycznych w metalu A mogą przejść do metalu B obsadzając tam niższe poziomy

energetyczne (ale znajdujące się powyżej poziomu Fermiego E_{FB}). Ten przepływ elektronów wywołuje powstawanie różnicy potencjałów między A i B i trwa aż do osiągnięcia wspólnego dla obu metali poziomu Fermiego (rys. 32.7b). W praktyce liczba swobodnych elektronów, które przeszły z metalu A do B jest małym ułamkiem ogólnej ich liczby, a zatem ostateczny efekt jest taki, jakby powierzchnie krzywych rozkładu energii w obu metalach pozostały niezmienione, a tylko uległy wzajemnemu przesunięciu objawiającemu się w postaci: 1) wyrównania poziomów Fermiego, 2) wytworzenia takiej różnicy potencjałów U_k między metalami, że iloczyn eU_k jest prawie równy różnicy prac wyjścia obu metali.



Rys. 32.7

Trzeba jednak podkreślić, że jeśli z 2 metali utworzymy obwód zamknięty, to na obu spojeniach powstaną różnice potencjałów U_k równe co do wartości, lecz przeciwne co do znaku, a zatem w obwodzie takim — jeśli oba spojenia są w tej samej temperaturze — wypadkowa siła elektromotoryczna (SEM) jest równa zeru.

Wystarczy jednak zmienić temperaturę jednego ze spojeń, by zmieniły się prace wyjścia i potencjał kontaktowy. W obwodzie wytworzy się wypadkowa siła elektromotoryczna, zwana *siłą termoelektryczną*, zdolna do wywołania przepływu prądu w obwodzie. Mówimy wtedy o *termoogniwach*. Przyrosty wartości siły termoelektrycznej są rzędu mikrowoltów na 1 stopień różnicy temperatur między spojeniami.

Badania wykazały, że siła termoelektryczna E_t dla większości par metali zmienia się według równania:

$$E_t = \alpha t + \tfrac{1}{2}\beta t^2,$$

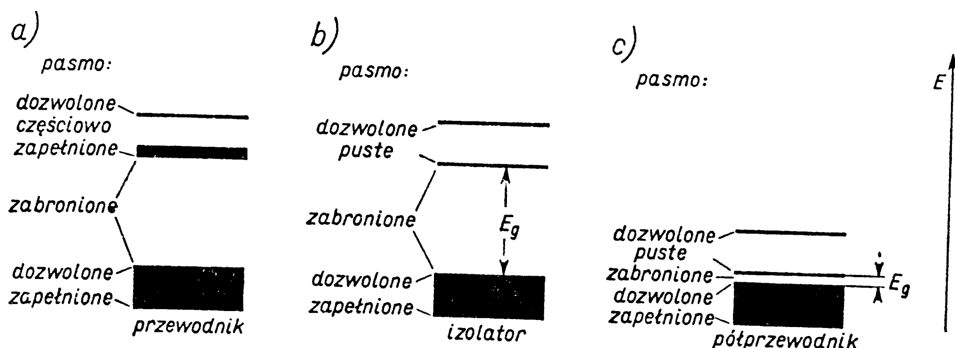
gdzie t oznacza różnicę temperatur między spojeniami, α i β — współczynniki stałe, charakterystyczne dla danej pary metali. Dla niektórych metali wartość β jest bliska zeru i wtedy w przybliżeniu siła termoelektryczna jest proporcjonalna do różnicy temperatur spojeń.

O zastosowaniu zjawiska termoelektrycznego mówiliśmy w § 13.4.

32.8. Model pasmowy ciała stałego

Kończąc omawianie właściwości metali powróćmy raz jeszcze do rys. 32.2. Widzimy na nim dyskretne poziomy energetyczne izolowanego atomu, rozciągnięte w coraz szersze dozwolone quasi-pasma energetyczne (ciemne) poprzedzielane pasmami zabronionymi (jasnymi).

Pasma energetyczne dozwolone położone niżej są w całości wypełnione elektronami. Tworzą one tzw. pasma zapełnione. Inaczej może się przedstawiać sprawa zapełnienia pasm energetycznych wyższych. Rozpatrzenie trzech odrębnych możliwości prowadzi do podziału ciał stałych na: przewodniki metaliczne, izolatory (dielektryki) i półprzewodniki.



Rys. 32.8

Przewodnik (metaliczny) charakteryzuje się tym, że najwyższe pasmo energetyczne zawierające elektrony walencyjne nie jest całkowicie obsadzone (rys. 32.8a). Wobec gęstego rozmieszczenia dozwolonych poziomów energetycznych w tym pasmie małe ilości dostarczanej energii wystarczają do przeniesienia elektronu na wyższy poziom energetyczny w obrębie tego samego pasma. Energii tej może dostarczyć na przykład słabe przyłożone pole elektryczne.

W izolatorach (dielektrykach) sprawa przedstawia się zupełnie inaczej. W izolatorach najwyższe pasmo energetyczne zawierające elektrony jest całkowicie zapełnione — co jest bardzo charakterystyczne — oddzielone jest od następnego dozwolonego pasma energetycznego szeroką ***przerwą energetyczną*** E_g (rys. 32.8b). Przerwa jest tak duża, że ani stosunkowo dość silne pole elektryczne, ani ogrzanie nie wystarczają do przeniesienia elektronu do wyższego pasma dozwolonego. W bardzo silnych polach może oczywiście nastąpić przejście pewnej liczby elektronów z pasma zapełnionego do wyższego pasma dozwolonego (tzw. przebicie dielektryka), co umożliwia przepływ prądu elektrycznego w izolatorze.

Trzecią grupę ciał stałych stanowią półprzewodniki. Schematyczne przedstawienie pasm energetycznych w półprzewodnikach widzimy na rys. 32.8c. Podobnie jak w przypadku izolatorów, najwyższe pasmo energetyczne zawierające elektrony, tak zwane ***pasmo walencyjne***, jest całkowicie zapełnione, ale tym razem przerwa energetyczna E_g

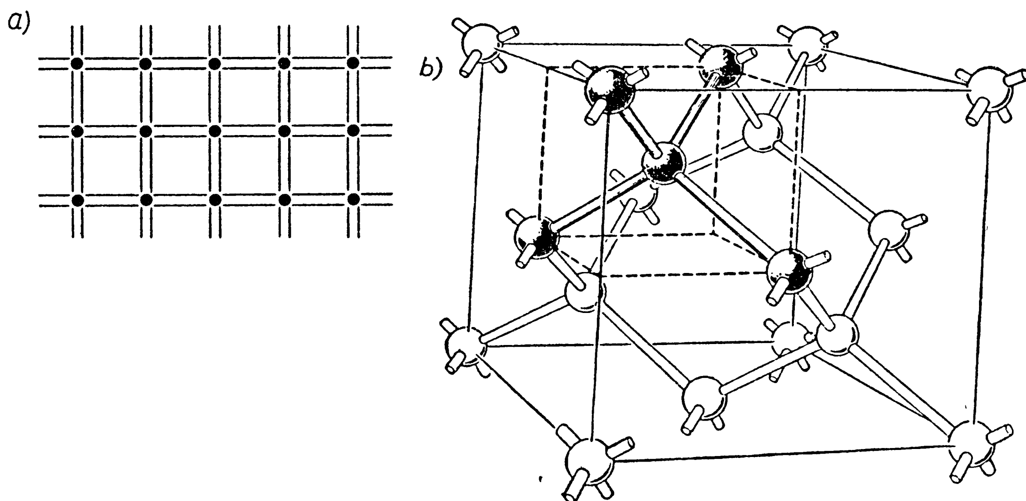
między ostatnim pasmem zapełnionym i następnym dozwolonym jest mała. Dzięki temu małe dawki energii związane np. z ogrzaniem, naświetleniem itp. wystarczają do przeniesienia elektronów z pasma walencyjnego do pasma wyższego, zwanego *pasmem przewodnictwa*.

32.9. Struktura elektronowa i krystaliczna krzemu i germanu. Wiązania kowalencyjne

Od uwag ogólnych dotyczących podziału ciał stałych na przewodniki, izolatory i półprzewodniki przejdziemy do bardziej szczegółowego omówienia tych ostatnich ze względu na ich stale wzrastające zastosowania.

Głównymi przedstawicielami półprzewodników są: krzem i german. Poza tym do grupy tej należą między innymi: diament, różne tlenki — jak np. Cu_2O, ZnO, siarczki — jak np. PbS, CdS, selen i selenki — jak np. $CdSe_2$, arsenki, tellurki itp.

Zajmiemy się bliżej krzemem i germanem. Krzem, pierwiastek o liczbie atomowej $Z = 14$, ma całkowicie wypełnione dwie powłoki elektronowe (razem 10 elektronów), a ponadto w powłoce (o $n = 3$) ma cztery elektrony o charakterystykach $3s^2$ $3p^2$. German, pierwiastek o liczbie atomowej $Z = 32$, ma całkowicie wypełnione trzy powłoki elektronowe (razem 28 elektronów), a ponadto w powłoce (o $n = 4$) ma cztery elektrony o charakterystykach $4s^2$ $4p^2$.



Rys. 32.9

Zarówno dla kryształów krzemu, germanu, jak i dla bardzo pod tym względem typowego diamentu (węgiel — 4 elektrony zewnętrzne $2s^2$ $2p^2$) charakterystyczne jest występowanie wiązań kowalencyjnych.

Wiązanie kowalencyjne jest wiązaniem dwuelektronowym, w którym biorą udział dwa elektrony walencyjne, należące równocześnie do dwóch atomów. Schematycznie

dla sieci dwuwymiarowej przedstawia to rys. 32.9a (czarne kółeczka — atomy, dwa równoległe odcinki — odpowiednik kowalencyjnego wiązania za pomocą pary elektronów).

Rysunek 32.9b przedstawia w analogiczny sposób trójwymiarową strukturę krystaliczną i komórkę elementarną jednakową dla diamentu, krzemu, germanu (zwaną strukturą krystaliczną diamentu). Jak widać, każdy z czterech elektronów walencyjnych każdego z atomów jest „zaangażowany" w wytworzenie wiązania walencyjnego między atomami, przy czym każda para wiążących elektronów jest dość sztywno związana z parą atomów, które łączy. A zatem brak jest w takiej strukturze swobodnych elektronów, stanowiących nośniki ładunku. Takie ciała powinny być doskonałymi izolatorami.

Jak zobaczymy dalej, sytuacja może się radykalnie zmienić pod wpływem: a) ogrzania, b) istnienia domieszek lub innych defektów sieci krystalicznej, c) naświetlenia.

Wyobraźmy sobie np., że taki kryształ silnie ogrzewamy. Wpływa to na zwiększenie amplitudy drgań atomów w sieci krystalicznej. Wreszcie drgania mogą stać się tak silne, że niektóre wiązania ulegną zerwaniu i zwolnione w ten sposób elektrony zaczną się zachowywać jak elektrony swobodne: przyłożenie zewnętrznego napięcia elektrycznego może wywołać ich ruch uporządkowany. W wysokiej temperaturze kryształy takie staną się przewodzące. Mamy więc do czynienia z ciałem nieprzewodzącym w pewnych warunkach, a przewodzącym w innych, co można uważać za jedno z uzasadnień nadawanej mu nazwy półprzewodnika.

32.10. Półprzewodniki samoistne

Jako podstawę wyjściową do rozważań teoretycznych, dotyczących właściwości półprzewodników, przyjmuje się model tak zwanego *półprzewodnika samoistnego*. Pod tą nazwą rozumie się kryształ idealny o wiązaniach kowalencyjnych, nie zawierający żadnych domieszek ani zakłóceń sieci krystalicznej.

Zgodnie z tym, co powiedzieliśmy w paragrafie poprzednim, półprzewodniki samoistne w temperaturze zera bezwzględnego zupełnie nie przewodzą elektryczności. Z punktu widzenia modelu pasmowego ciała stałego odpowiada to całkowitemu zapełnieniu energetycznego pasma walencyjnego. Gdyby to pasmo nie było całkowicie zapełnione, mielibyśmy znowu taką sytuację, jak w przypadku przewodników metalicznych.

Wystąpienie przewodnictwa elektrycznego w półprzewodnikach samoistnych możliwe jest po dostarczeniu odpowiedniej porcji energii — np. przez ogrzanie. Ta porcja energii musi wystarczyć na przeniesienie elektronu z zapełnionego pasma walencyjnego do najbliższego dozwolonego pasma przewodnictwa, oddalonego od pasma walencyjnego o przerwę energetyczną E_g (por. tabelę 32.2). Przy tym przeniesieniu elektronu ulega zerwaniu wiązanie kowalencyjne, pojawia się swobodny elektron w pasmie przewodnictwa, a w miejscu zwolnionym przez elektron pojawia się „dziura". A więc w temperaturze odpowiednio wysokiej (o jej wysokości decyduje przerwa energetyczna E_g danego materiału) wytwarzają się w półprzewodniku samoistnym dwa częściowo

zapełnione pasma: pasmo przewodnictwa z pewną liczbą swobodnych elektronów i pasmo walencyjne z identyczną liczbą zwolnionych miejsc — dziur. W takiej sytuacji możliwe już jest przewodzenie elektryczności.

Przeanalizujemy dokładniej przewodnictwo elektryczne półprzewodników samoistnych. Ma ono charakter złożony, swój udział bowiem wnoszą zarówno elektrony, poruszające się w pasmie przewodnictwa, jak i dziury — w pasmie walencyjnym. W związku z tym rozróżniamy przewodnictwo elektronowe i przewodnictwo dziurowe. *Przewodnictwo elektronowe półprzewodników* ma charakter dokładnie taki sam, jak przewodnictwo elektronowe w metalach.

Tabela 32.2

Przerwy energetyczne E_g kilku typowych półprzewodników w temperaturze pokojowej

Substancja	E_g eV
diament	6,0
german	0,68–0,72
krzem	1,1
selenek kadmu	1,74
selenek ołowiu	0,27
siarczek kadmu	2,42
siarczek ołowiu	0,34–0,37
tellurek kadmu	1,45

Przewodnictwo dziurowe wymaga bliższego rozpatrzenia. Dziura wytworzona w pewnym miejscu sieci krystalicznej nie utrzymuje się tam trwale. Pod wpływem przyłożonego elektrycznego pola przesuwa się na jej miejsce elektron z sąsiedniego wiązania kowalencyjnego zamykając wiązanie zerwane (wypełniając dziurę), lecz równocześnie pozostawiając nową dziurę w swoim poprzednim położeniu. Nowa dziura zostaje z kolei zapełniona przez następny sąsiedni elektron i w ten sposób faktycznemu ruchowi elektronów od wiązania do wiązania towarzyszy tak zwana migracja dziur (w przeciwnym kierunku) w pasmie walencyjnym. Mamy więc znów do czynienia z ruchem nośników ładunków — z wkładem wnoszonym do prądu przewodzonego przez półprzewodnik. *Zasadnicza różnica między przewodnictwem elektronowym i dziurowym* polega na tym, że o ile w przewodnictwie elektronowym można mówić o ruchu swobodnych elektronów przebywających bez przeszkód długie drogi swobodne wewnątrz sieci krystalicznej, o tyle w przewodnictwie dziurowym, w którym elektron przesuwa się tylko od danego wiązania do sąsiedniego nie ma już mowy o ruchu elektronów swobodnych.

Jak już wspominaliśmy, ruch dziur w polu elektrycznym odbywa się w kierunku przeciwnym do ruchu elektronów: można więc uważać, że dziura zachowuje się jak ładunek dodatni równy co do wartości ładunkowi elektronu.

Na całkowitą gęstość j_t* prądu w półprzewodniku samoistnym składa się gęstość j_e prądu elektronowego i gęstość j_h* prądu dziurowego

$$j_t = j_e + j_h.$$

Ale z teorii przewodnictwa elektronowego metali wiemy, że spełnione są następujące zależności:

$$j = \gamma E \quad \text{i} \quad \gamma = en\mu,$$

gdzie γ oznacza przewodność właściwą, E — natężenie pola elektrycznego, e, μ, n — odpowiednio ładunek, ruchliwość i koncentrację elektronów.

Stosując analogiczne zależności do przypadku półprzewodnika samoistnego otrzymujemy następujące wyrażenie na całkowitą przewodność właściwą γ_t półprzewodnika:

$$\gamma_t = \gamma_e + \gamma_h,$$

gdzie γ_e i γ_h oznaczają odpowiednio przewodność właściwą elektronową i dziurową

$$\gamma_t = |e|(n_e\mu_e + n_h\mu_h).$$

(Obie ruchliwości μ_e i μ_h przyjęto jako dodatnie; stąd we wzorze wartość bezwzględna ładunku elementarnego charakterystycznego dla obu rodzajów nośników prądu: elektronów i dziur.)

W półprzewodniku samoistnym koncentracja n_e elektronów w pasmie przewodnictwa i koncentracja n_h dziur w pasmie walencyjnym jest jednakowa. Ruchliwości obu nośników prądu jednak niewątpliwie się różnią w związku z różnym mechanizmem przewodzenia elektronowego i dziurowego. Ruchliwość elektronów znacznie przewyższa ruchliwość dziur, jak to widać z zestawienia kilku przybliżonych wartości μ_e i μ_h.

Rodzaj substancji	μ_e $\frac{m^2}{V \cdot s}$	μ_h $\frac{m^2}{V \cdot s}$
diament	0,18	0,12
german	0,38	0,18
krzem	0,16	0,04

Ruchliwości obu rodzajów nośników dla danego rodzaju substancji są stałe w szerokich granicach stosowanych pól elektrycznych. Na przykład dla germanu można przyjmować, że μ_e jest stałe do wartości $9 \cdot 10^4$ V/m, a μ_h do $10 \cdot 10^4$ V/m.

Z różnic ruchliwości obu nośników wynika, że większy jest udział prądu elektronowego niż prądu dziurowego w prądzie całkowitym.

32.11. Półprzewodniki domieszkowe

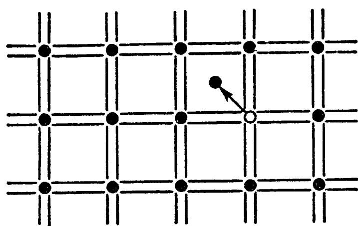
Właściwości półprzewodników zależą w bardzo dużym stopniu od rodzaju i liczby atomów domieszek wprowadzonych do kryształów półprzewodnikowych. Wpływ domieszek zaznacza się już przy bardzo małych ich ilościach, rzędu 1 atom domieszki

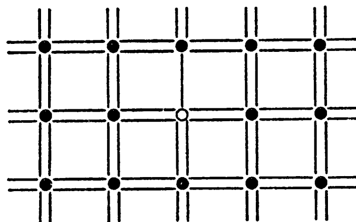
* Po angielsku *total* — całkowity, *hole* — dziura.

na 10^7–10^8 atomów półprzewodnika. Na przykład wprowadzenie boru do krzemu w ilości 1 atom boru na 10^5 atomów krzemu powoduje już tysiąckrotny wzrost przewodności krzemu.

Przeanalizujemy kolejno rolę domieszek atomów z V i III grupy układu okresowego w kryształach germanu lub krzemu.

Wprowadzenie do sieci krystalicznej germanu (krzemu, diamentu) atomu pierwiastka pięciowartościowego, np. fosforu, arsenu, antymonu zmienia sytuację przedstawioną na rys. 32.9a na sytuację przedstawioną na rys. 32.10. Cztery elektrony walencyjne atomu domieszki zastępują z powodzeniem cztery elektrony walencyjne właściwego atomu sieci tworząc wiązania kowalencyjne. Pozostaje jednak elektron piąty, słabo związany z jądrem swego atomu, a więc taki, który nawet w temperaturze niższej od pokojowej może być łatwo oderwany dzięki drganiom termicznym sieci. Zachowuje się on już wtedy jak elektron swobodny, biorący udział w przewodnictwie elektronowym kryształu. Takie domieszki dostarczające dodatkowych elektronów przewodnictwa noszą nazwę *domieszek donorowych**: atom domieszki w postaci fosforu, arsenu, antymonu jest *donorem*.


Rys. 32.10


Rys. 32.11

Półprzewodnik domieszkowy z donorami nosi nazwę *półprzewodnika typu n*. Sieć krystaliczna takiego półprzewodnika różni się od normalnej sieci krystalicznej germanu lub krzemu tym, że w niektórych węzłach sieci znajdują się dodatnie jony domieszki sztywno związane z resztą sieci.

Inaczej przedstawia się sytuacja (rys. 32.11) po wprowadzeniu do półprzewodnika atomu trójwartościowej domieszki, na przykład atomu boru, glinu, indu, galu. Tym razem do wytworzenia doskonałej sieci z wiązaniami kowalencyjnymi brak jednego walencyjnego elektronu. Jeśli temperatura jest dostatecznie wysoka, to elektrony z sąsiednich wiązań wykazują tendencję do wypełniania tej luki w wiązaniu sieci, a zatem defekt sieci wywołany wprowadzeniem domieszki trójwartościowej oddziałuje na otoczenie jak dziura w półprzewodniku samoistnym. Tym razem przewodnictwo domieszkowe ma charakter dziurowy.

Półprzewodnik domieszkowy tego rodzaju nosi nazwę *półprzewodnika typu p***. Atomy domieszki przyłączają tym razem jeden elektron z pasma walencyjnego półprzewodnika; w związku z tym nadano tym atomom nazwę *akceptorów*. Powstające

* Po angielsku *negative* — ujemny, *donate* — dostarczać.

** Po angielsku *positive* — dodatni.

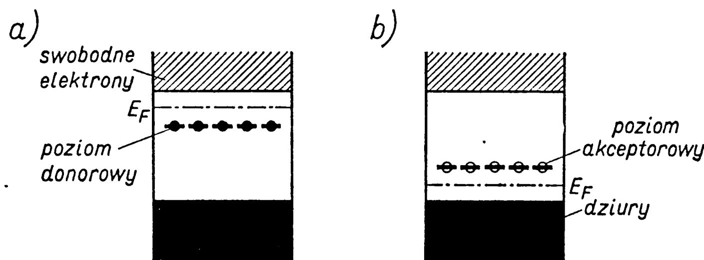
w opisany sposób jony ujemne domieszki nie biorą udziału w przewodnictwie półprzewodnikowym, gdyż są unieruchomione w określonym węźle sieci.

Rozwinięcie teorii półprzewodników domieszkowych ma duże znaczenie, gdyż praktycznie biorąc wszystkie półprzewodniki są półprzewodnikami nie samoistnymi, lecz domieszkowymi. Zwykle występują łącznie domieszki akceptorowe i donorowe. Półprzewodnik jest typu *n*, gdy przeważają domieszki donorowe, typu *p* — gdy przeważają domieszki akceptorowe. Właściwości półprzewodnika samoistnego wykazuje półprzewodnik z równą liczbą donorów i akceptorów.

32.12. Model pasmowy półprzewodnika domieszkowego

Inna interpretacja właściwości półprzewodników domieszkowych opiera się na modelu pasmowym.

Zacznijmy od półprzewodnika typu *n*. Nieliczne stosunkowo atomy domieszki z grupy V układu okresowego wprowadzają, jak już mówiliśmy, swój piąty elektron walencyjny jako defekt sieci. Elektron ten znajduje się na dyskretnym poziomie energetycznym (inne atomy domieszki są daleko i nie powodują rozszczepienia poziomu energetycznego w pasmo). Jest on słabo związany ze swoim jądrem atomowym, a więc dla niego energia jonizacji może być znacznie mniejsza niż energia jonizacji dla któregokolwiek elektronu uczestniczącego w wiązaniu kowalencyjnym. A zatem na rysunku 32.12a, przedstawiającym pasmo walencyjne (ciemne) i najbliższe dozwolone (na razie puste — lekko zakreskowane na rysunku) pasmo przewodnictwa półprzewodnika typu *n*, pojawia się w przerwie energetycznej E_g dyskretny energetyczny poziom domieszki donorowej (gruba kreskowana linia na rysunku).



Rys. 32.12

Ten poziom donorowy jest całkowicie obsadzony w temperaturze zera bezwzględnego (wszystkie elektrony walencyjne donorów są związane ze swymi atomami). Przypominają o tym czarne kółka zaznaczone na poziomie donorowym. W temperaturach wyższych sytuacja przedstawia się inaczej. Wobec małej odległości (w skali energii) poziomu donorowego od dolnej granicy pasma przewodnictwa mała porcja energii, wynikająca np. z ogrzania, wystarcza do przeniesienia elektronu domieszki do pasma przewodnictwa. W ten sposób część poziomów w pasmie przewodnictwa zostaje obsadzona elektronami, które (jako swobodne) po przyłożeniu zewnętrznego pola uczestniczą w przewodnictwie elektronowym półprzewodnika.

Warto podkreślić, że w półprzewodniku *n* nie powstają dziury w pasmie walencyjnym. Przewodnictwo jest tym razem czysto elektronowe. Stanowi to zasadniczą różnicę w stosunku do półprzewodnika samoistnego, w którym liczba dziur w pasmie walencyjnym i swobodnych elektronów w pasmie przewodnictwa jest jednakowa.

W analogiczny sposób przedstawimy model pasmowy półprzewodnika typu *p*. Tym razem — jak pamiętamy — brakuje w atomie domieszki (z III grupy układu okresowego) jednego elektronu do wytworzenia czterech wiązań kowalencyjnych. Mała porcja energii (wynikająca np. z ogrzania) wystarcza do tego, aby elektron z pasma walencyjnego półprzewodnika przeszedł na dozwolony „akceptorowy" dyskretny poziom energetyczny atomu domieszki, leżący bardzo blisko górnego poziomu energetycznego pasma walencyjnego. Ten poziom akceptorowy zaznaczony jest również grubą kreskowaną linią na rys. 32.12b. Białe kółeczka zaznaczone na tym poziomie przypominają, że w temperaturze zera bezwzględnego poziom ten jest całkowicie nieobsadzony. Jednak w temperaturze pokojowej sytuacja przedstawia się inaczej, każdy z atomów akceptorowej domieszki dołącza jeden elektron z pasma walencyjnego półprzewodnika. W pasmie tym wytwarza się zatem odpowiednia liczba dziur. Tym razem nie ma w ogóle elektronów w paśmie przewodnictwa: półprzewodnik typu *p* wykazuje wyłącznie przewodnictwo dziurowe. W węzłach sieci pojawiają się ujemne jony domieszki.

Spoglądając uważnie na dwa ostatnie rysunki znajdziemy na nich dodatkowe poziomy złożone z dłuższych i krótszych kresek, oznaczone znanym już symbolem E_F, dotyczącym poziomu Fermiego. Ale dotychczas zdefiniowaliśmy poziom Fermiego tylko w odniesieniu do metali. Definicja tego poziomu odniesiona do półprzewodników (lub izolatorów) jest bardziej skomplikowana. Wymaga ona wybrania dwóch najbliższych takich poziomów energetycznych E_1 i E_2, z których jeden wykazuje w stosunku do maksymalnych możliwości obsadzenia go elektronami niedobór wynoszący np. X%, a drugi wykazuje właśnie obsadzenie w X%. Średnia arytmetyczna energii obu poziomów przedstawia, zgodnie z przyjętą umową, energię Fermiego, a odpowiedni poziom Fermiego leży w środku odległości między E_1 i E_2.

Przeanalizujmy położenie poziomu Fermiego dla trzech poznanych już rodzajów półprzewodników.

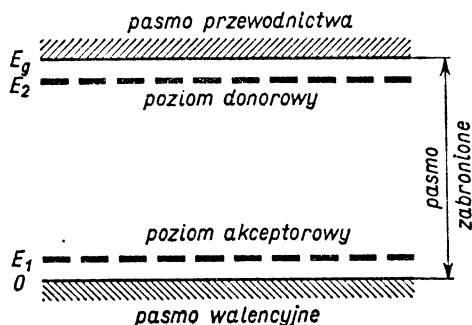
W półprzewodniku samoistnym poziom odpowiadający górnej granicy pasma walencyjnego jest w 100% obsadzony, a poziom odpowiadający dolnej granicy pasma przewodnictwa jest w 100% nie obsadzony w temperaturze zera bezwzględnego. A zatem poziom Fermiego w półprzewodniku samoistnym leży dokładnie w środku przerwy energetycznej E_g. Sytuacja taka obowiązuje zarówno dla zera bezwzględnego, jak i dla wszelkich innych temperatur.

W półprzewodniku typu *n* (w temperaturze TK) najbliższymi poziomami energetycznymi, w których braki i nadwyżki elektronów mogą być porównywane, są: poziom donorowy i poziom bardzo bliski dolnej granicy pasma przewodnictwa. Poziom Fermiego leży więc mniej więcej w połowie odległości wspomnianych poziomów, ogólnie biorąc stosunkowo blisko dolnej granicy pasma przewodnictwa.

W półprzewodniku typu *p* (w temperaturze TK) nadwyżki występują w poziomie

akceptorowym, braki zaś w pobliżu górnej granicy pasma walencyjnego (tam, gdzie pojawiły się dziury). A zatem poziom Fermiego przypada mniej więcej w połowie odległości wspomnianych poziomów. Tym razem leży on stosunkowo blisko górnej granicy pasma walencyjnego.

Jak już mówiliśmy w § 32.11, mamy zwykle do czynienia z półprzewodnikami zawierającymi równocześnie domieszki donorowe i akceptorowe. W takim przypadku charakterystyczny wykres poziomów energetycznych ma postać przedstawioną na rys.



Rys. 32.13

32.13. Tym razem jako zerowy poziom odniesienia przyjęta jest górna granica pasma walencyjnego. Dolna granica pasma przewodnictwa jest odległa o E_g, poziom akceptorowy odpowiada E_1, poziom donorowy — E_2.

Do otrzymania swobodnego elektronu z domieszki donorowej potrzebna jest energia aktywacji $E_g—E_2$, a do wytworzenia dziury przy udziale domieszki akceptorowej — energia E_1. Przykładowe wartości tych energii dla germanu i krzemu wynoszą odpowiednio:

	Ge	Si
dla fosforu jako donora	0,012 eV	i 0,045 eV,
dla boru jako akceptora	0,0104 eV	i 0,045 eV.

Są więc one porównywalne z energią kT odpowiadającą temperaturze pokojowej, równą 0,026 eV. Nic więc dziwnego, że występuje duży wpływ jonizacji termicznej donorów i akceptorów na przewodnictwo elektryczne germanu i krzemu nawet w temperaturze pokojowej.

W warunkach równowagi termicznej, gdy badany półprzewodnik nie jest poddany żadnym oddziaływaniom zewnętrznym, istnieje w nim określona koncentracja elektronów i dziur charakterystyczna dla danej temperatury T.

Zmianę wspomnianej koncentracji nośników w danej temperaturze można wywołać przez naświetlanie odpowiednim promieniowaniem, bombardowaniem szybkimi, naładowanymi cząstkami lub przez „wstrzykiwanie" nośników do badanej próbki przez kontakt metaliczny. Usunięcie czynnika zakłócającego powoduje po pewnym czasie powrót do równowagi dzięki zjawisku rekombinacji nadmiaru elektronów i dziur.

Na zakończenie jeszcze kilka uwag ogólnych.

Nasuwa się pytanie, dlaczego tyle uwagi poświęciliśmy domieszkom z grupy V i III układu okresowego pierwiastków, skoro w półprzewodnikach mogą istnieć domieszki także innych rodzajów. Wiadomo np., że często spotykane są domieszki w postaci atomów metali dwuwartościowych, takich jak cynk lub jednowartościowych, takich jak miedź.

Domieszki w postaci atomów wymienionych przykładowo metali wytwarzają swoje poziomy (akceptorowe) nawet w większej ilości niż akceptory z grupy III, gdyż cynk wytwarza 2 poziomy akceptorowe, a miedź aż trzy. Wszystkie one jednak leżą znacznie dalej (w skali energii) od górnego poziomu pasma walencyjnego, a więc charakterystyczne dla nich wartości energii E_1 (rys. 32.13) są większe niż odpowiednie wartości E_1 dla poziomów akceptorowych domieszek z grupy III. Właśnie to decyduje o mniejszym wpływie domieszek z innych grup na proces generowania nośników prądu. Pojawienie się nowych dozwolonych poziomów energetycznych w części środkowej przerwy energetycznej E_g jest nawet wyraźnie niekorzystne, gdyż zwiększa możliwości rekombinacji nadmiaru elektronów i dziur. Badania wykazują, że istnienie poziomu domieszkowego gdzieś w pobliżu $E_g/2$ ogromnie zwiększa prawdopodobieństwo wystąpienia procesu rekombinacji rozbijając go na dwa: przejście elektronu z pasma przewodnictwa na poziom atomu domieszki i „zderzenie" dziury z tym atomem domieszki. Kilka liczb wykaże wpływ domieszek na czas rekombinacji. Dla idealnie czystego kryształu germanu teoretycznie obliczony czas rekombinacji wynosi około 1 sekundy, w najczystszych otrzymywanych w laboratoriach kryształach ten czas wynosi około 0,005 s, a domieszki miedzi w ilości $1:10^8$ atomów redukują ten czas do 10^{-5} s.

W wielu urządzeniach półprzewodnikowych ważne jest dla właściwego ich funkcjonowania, aby czas rekombinacji był dość długi. W związku z tym opracowane są różne metody usuwania z półprzewodników niepożądanych domieszek.

32.13. Zależność przewodności elektrycznej półprzewodników od temperatury

Jednym z najważniejszych czynników decydujących o przewodności półprzewodników jest temperatura. W półprzewodnikach samoistnych — jak już wiemy — w temperaturze zera bezwzględnego nie ma w ogóle elektronów przewodnictwa. Pojawiają się one w miarę wzrostu temperatury, gdyż niektóre elektrony z pasma walencyjnego otrzymują energię wystarczającą do przejścia przerwy energetycznej E_g. Można wykazać, że półprzewodnik samoistny w temperaturze T K ma przewodność elektryczną właściwą γ zależną wykładniczo od temperatury:

$$\gamma = \gamma_0 e^{-E_g/2kT}$$

gdzie γ_0 oznacza współczynnik o wartości 10^5 dla większości półprzewodników, k — stałą Boltzmanna.

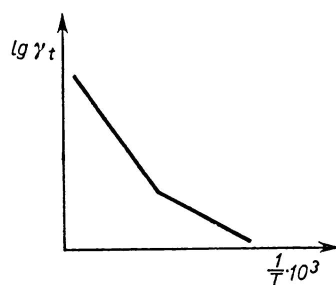
W praktyce jednak mamy do czynienia z półprzewodnikami domieszkowymi, dla których charakterystyczne są energie aktywacji o wartościach E_1 lub $(E_g - E_2)$ (por. rys. 32.13). Oznaczmy ogólnie energię aktywacji domieszki donorowej lub akceptorowej przez E_a. Wtedy — dla półprzewodnika domieszkowego — wyrażenie na całko-

witą przewodność elektryczną właściwą γ_t napiszemy w postaci

$$\gamma_t = \gamma_0 e^{-E_g/2kt} + \gamma_1 e^{-E_a/2kT},$$

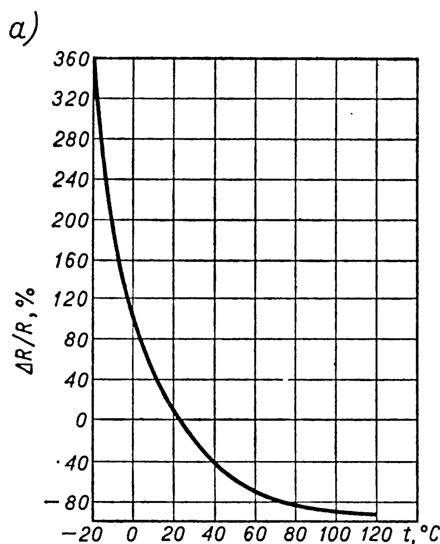
gdzie składnik pierwszy dotyczy przewodności samoistnej, składnik drugi — przewodności domieszkowej. Składnik pierwszy jest bardzo mały w niskich temperaturach, w tych warunkach o przewodności decydują domieszki. Wraz ze wzrostem temperatury stopniowo coraz większą rolę zaczyna odgrywać właśnie przewodność samoistna. Przy pewnej temperaturze kończy się wpływ atomów domieszek na zjawisko przewodzenia prądu. W temperaturach większych od wspomnianej mamy do czynienia wyłącznie z przewodnością samoistną półprzewodnika.

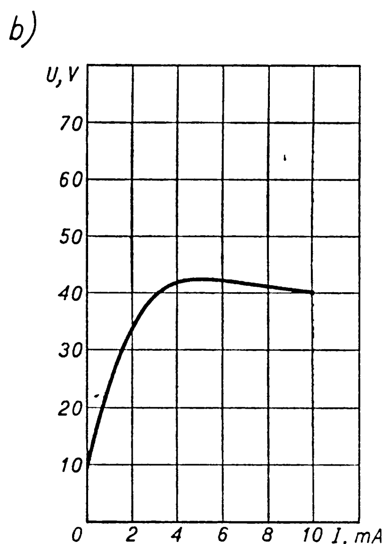
Uproszczony przebieg zależności $\lg \gamma_t$ od $1/T$ dla większości półprzewodników domieszkowych ma kształt podobny do przedstawionego na rys. 32.14. Jest to linia łamana, której lewy odcinek odpowiada wysokim temperaturom, charakteryzuje zatem przewodność samoistną, natomiast prawy odcinek odpowiada temperaturom niskim, charakteryzuje zatem przewodność domieszkową. Współczynniki kierunkowe obu liniowych części temperaturowych charakterystyk półprzewodnika pozwalają na wyznaczanie wartości E_g i E_a, a więc bardzo ważnych parametrów charakteryzujących półprzewodnik.



Rys. 32.14

Termistory. Nazwę termistorów nadaje się urządzeniom oporowym półprzewodnikowym, zbudowanym z materiałów o dużym współczynniku temperaturowym oporu.





Rys. 32.15

Termistory mogą być wykonane z dowolnego rodzaju półprzewodników. W zależności od ich przeznaczenia dobiera się odpowiednio rodzaj materiału uwzględniając przede wszystkim takie parametry jak współczynnik temperaturowy oporu, temperaturę topnienia, odporność chemiczną substancji, stabilność działania.

Każdy termistor — o dowolnym kształcie geometrycznym: płytki, kulki, nici itp. — ma dwie przeciwległe elektrody, za pomocą których włączany jest do obwodu elektrycznego. Opór termistora jest znacznie większy od oporu omowego reszty obwodu i zależy silnie od temperatury. A zatem praktycznie o natężeniu prądu płynącego w obwodzie elektrycznym decyduje rodzaj włączonego termistora i jego temperatura, przy czym ze wzrostem temperatury prąd rośnie, a w miarę obniżania temperatury — prąd maleje. Opór termistora może się zmieniać od 10 do 1000 razy, zależnie od rodzaju użytego półprzewodnika (w tym samym przedziale temperatur). Przebieg zależności zmian oporu od temperatury dla określonego termistora podany jest przykładowo na rys. 32.15a, a charakterystyka prądowo-napięciowa (czyli zależność między napięciem i prądem) na rys. 32.15b.

Zastosowania termistorów są w tej chwili bardzo liczne. Niektóre z nich wymienialiśmy już w poprzednich rozdziałach, jak np. termometry termistorowe, manometry próżniowe, o działaniu opartym na właściwościach termistorów, bolometry półprzewodnikowe. Poza tym termistory znalazły zastosowania między innymi jako regulatory prądu, stabilizatory napięcia, przekaźniki czasu itp.

32.14. Fotoprzewodnictwo półprzewodników

W dotychczasowych naszych rozważaniach zakładaliśmy, że badany półprzewodnik nie jest poddany działaniu światła. Prąd elektryczny płynący w obwodzie w tych warunkach nazywamy *prądem ciemnym* i opór, jaki w tych warunkach ma półprzewodnik, nazywamy również *oporem ciemnym*.

Oświetlenie powierzchni półprzewodnika odpowiednio dobranym rodzajem promieniowania powoduje skokowy wzrost prądu ciemnego. Różnica prądu przy zastosowaniu światła i prądu ciemnego daje tak zwany fotoprąd I_f. Zjawisko opisane nosi nazwę fotoprzewodnictwa, a półprzewodnik, w którym zachodzą powyższe zmiany — fotoopornika.

Nietrudno zrozumieć istotę fotoprzewodnictwa. Kwanty promieniowania $h\nu$, mające energię większą lub co najmniej równą energii aktywacji elektronów domieszek donorowych, mogą wywołać pojawienie się dodatkowych nośników prądu. Opór próbki wskutek tego maleje, a przewodność i prąd rosną w stosunku do wartości, jakie występowały w ciemności w tej samej temperaturze. Jeśli zastosujemy takie promieniowanie krótkofalowe, że energia kwantu jest równa lub nieco większa od E_g, występuje przenoszenie elektronów z pasma walencyjnego do pasma przewodnictwa. W tych warunkach generowana jest równocześnie para nośników prądu: elektron i dziura. Badania wykazały, że zastosowanie światła o kwantach znacznie większych od E_g, czyli światła bardziej krótkofalowego, daje znacznie słabszy efekt z powodu silnej jego absorpcji w krysztale półprzewodnika.

Dla półprzewodników samoistnych (bez domieszek) istnieje oczywiście wyraźna częstotliwość progowa ν_0 określająca wystąpienie fotoprzewodnictwa, spełniająca zależność

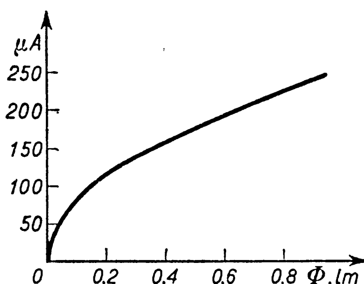
$$h\nu_0 = E_g.$$

W półprzewodnikach typu n częstotliwość progowa jest znacznie mniejsza wobec mniejszej wartości energii aktywacji (E_g—E_2).

Dla różnych półprzewodników częstotliwości progowe mają bardzo różne wartości z zakresu nadfioletu, światła widzialnego i podczerwieni.

Liczba elektronów fotoprzewodnictwa może być bardzo duża. Na przykład w siarczku kadmu (CdS) osiągnięto przy odpowiednim oświetleniu koncentrację fotoelektronów przewyższającą 10^4 razy koncentrację elektronów przewodnictwa ciemnego w tej samej temperaturze. Przeciętnie w półprzewodnikach przy niezbyt silnym oświetleniu fotoelektrony stanowią 20–30% ogólnej liczby elektronów przewodnictwa.

Warto może raz jeszcze podkreślić, że w opisywanym zjawisku, które można nazwać *zjawiskiem fotoelektrycznym wewnętrznvm*, energia promieniowania dostarczona z zewnątrz zostaje zmagazynowana wewnątrz fotoopornika — w sposób zewnętrznie niedostrzegalny — w postaci przeniesienia części elektronów na wyższy poziom energetyczny należący do pasma przewodnictwa.



Rys. 32.16

Fotooporniki. Mówiąc najogólniej, fotoopornik jest przyrządem półprzewodnikowym o przewodności elektrycznej zmieniającej się pod wpływem światła.

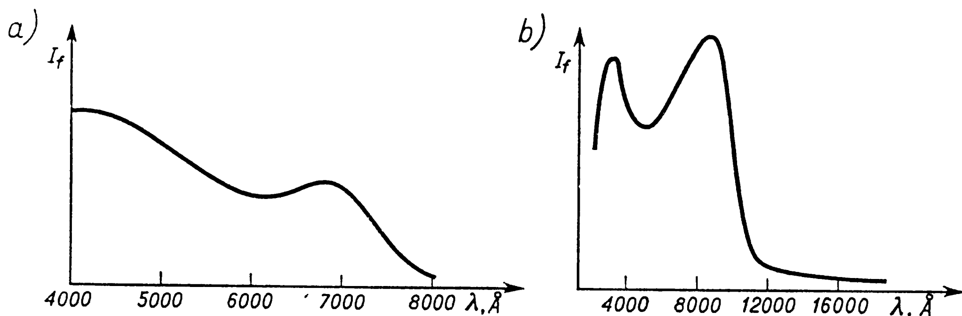
Najprostszy fotoopornik ma postać płytki (cienkiej warstwy) półprzewodnikowej o nałożonych na dwóch przeciwległych końcach elektrodach metalowych, zapewniających dobry kontakt metaliczny. Do urządzenia takiego dołącza się źródło napięcia i miernik prądu.

Dopóki fotoopornik znajduje się w ciemności, miernik wykazuje prąd ciemny. Zastosowanie oświetlenia powoduje wzrost prądu, przy czym zależność prądu od strumienia świetlnego (przy stałym napięciu) przebiega nieliniowo, jak to widać z rys. 32.16.

Na każdą zmianę w oświetleniu fotoopornik reaguje zmianą przewodności, a więc i zmianą prądu w obwodzie. Reakcja jest w zasadzie podobna do tej, którą wykazują obwody z fotokomórkami opartymi na zjawisku fotoelektrycznym zewnętrznym. Są jednak również znaczne różnice. Między innymi, czułość fotoopornika jest znacznie większa od czułości fotokomórek. O ile jednak reakcja fotokomórki jest praktycznie

natychmiastowa, o tyle większość fotooporników wykazuje pewną bezwładność, czyli opóźnienie reakcji przyrządu.

W praktycznych zastosowaniach fotooporników poważną ich wadę stanowi wspomniana już nieliniowa zależność I od Φ. Dzięki stałemu postępowi w dziedzinie półprzewodników możliwe jest jednak otrzymywanie coraz to nowych materiałów o coraz dogodniejszych parametrach, co przyczynia się do rozszerzenia zakresu praktycznych zastosowań fotooporników służących w zasadzie, podobnie jak i fotokomórki, do wykrywania i pomiarów energii świetlnej.



Rys. 32.17

O odpowiednim wyborze rodzaju fotoopornika do różnych celów decyduje w dużym stopniu przebieg jego charakterystyki spektralnej, to znaczy krzywej zależności fotoprądu od długości fali (lub częstotliwości) zastosowanego promieniowania. Przykłady różnego przebiegu takich charakterystyk przedstawione są na rys. 32.17a i b. Rys. 32.17a dotyczy fotoopornika o jednakowej w przybliżeniu czułości w dziedzinie widzialnej, a rys. 32.17b odpowiada fotoopornikowi (siarczek talu) o dużej czułości w dziedzinie bliskiej podczerwieni. Dużą czułość w dalszej podczerwieni (około 20 000 Å) wykazują fotooporniki z siarczku ołowiu. Fotooporniki o dużej czułości na podczerwień stosowane są w badaniach astrofizycznych, służą również do wykrywania z dużych odległości samolotów lub pocisków rakietowych.

Bardzo cennymi fotoopornikami dla badań w dziedzinie widzialnej są fotooporniki z siarczku kadmu o maksymalnej czułości spektralnej dla około 5500 Å. W fotooporniku tym osiągnięto bardzo wysoki stosunek prądu ciemnego do fotoprądu równy 1 : 150 000.

Zakończymy ten przegląd uwagą ogólną: największą fotoczułość wykazują te materiały półprzewodnikowe, które w ciemności zachowują się jak izolatory, a więc odznaczają się stosunkowo dużą przerwą energetyczną E_g. Do grupy tej należą: wymieniony już CdS (światło zielone), CdSe (światło czerwone) i CdTe (podczerwień).

32.15. Złącze *p-n*. Uwagi ogólne

Zasadniczą cechą omawianych dotychczas urządzeń półprzewodnikowych było to, że wykonane one były z półprzewodnika jednorodnego w całej objętości. Obecnie przejdziemy do omówienia innych urządzeń, może ważniejszych jeszcze z punktu wi-

dzenia zastosowań, wytwarzanych z jednego monokryształu, z tym jednak, że w pewnym przekroju zmienia się w sposób nagły półprzewodnik typu *p* na półprzewodnik typu *n*. Właśnie to miejsce nagłej zmiany półprzewodnika typu *p* w półprzewodnik typu *n* nazywa się ***złączem p-n***.

Prawie wszystkie urządzenia półprzewodnikowe ze złączami *p-n* otrzymywane są z kryształów germanu i krzemu. Jedna z metod ich wytwarzania np. w germanie polega na zmianie warunków hodowania kryształu przez wprowadzenie w pewnej chwili do stopionej masy germanu z mikroskopijną zawartością np. indu, z której wykrystalizowuje półprzewodnik typu *p*, nadmiaru domieszki donorowej (np. arsenu). Od tej chwili wykrystalizowuje półprzewodnik typu *n*. W jednym monokrysztale mamy obok siebie półprzewodniki typu *p* i *n*, rozdzielone złączem *p-n*.

Okazuje się, że urządzenia półprzewodnikowe z jednym lub kilkoma złączami *p-n* i odpowiednią liczbą elektrod mogą mieć różnorodne zastosowania, między innymi przede wszystkim jako diody krystaliczne (prostowniki) i triody krystaliczne, czyli tranzystory (wzmacniacze). Właśnie tym zastosowaniom poświęcimy następne paragrafy. Rozpoczniemy wyjaśnienie działania tych przyrządów od kilku uwag ogólnych, odnoszących się do złącza *p-n* w warunkach równowagi termicznej, bez zewnętrznego pola.

W chwili wytworzenia złącza *p-n* (bez przyłożonego zewnętrznego pola) mamy większą koncentrację elektronów przewodnictwa w obszarze *n* i większą koncentrację dziur w obszarze *p*. Takie warunki sprzyjają procesowi dyfuzji: elektrony dyfundują w stronę obszaru *p* pozostawiając za sobą dodatnio naładowane jony donorowe, a dziury dyfundują w stronę obszaru *n* pozostawiając za sobą ujemne jony akceptorowe. A zatem ta dwukierunkowa dyfuzja ładunków o przeciwnych znakach prowadzi do wytworzenia na złączu podwójnej warstwy elektrostatycznej. Pole elektryczne, które towarzyszy wytworzeniu wspomnianej warstwy, ma kierunek taki, że przeciwdziała dalszej dyfuzji ładunków przez złącze. Na złączu powstaje zatem różnica potencjałów, do której jeszcze powrócimy.

W warunkach równowagi termicznej, bez zewnętrznego pola, na złączu *p-n* stale odbywają się dwa ruchy nośników w przeciwnych kierunkach, czyli przepływają dwa prądy. Jeden z tych prądów, tak zwany ***prąd rekombinacji*** I_r, wiąże się z przepływem małej liczby dziur z obszaru *p* do *n*, gdzie ulegają one rekombinacji. Drugi prąd, tzw. ***prąd termiczny*** I_t, wynika z generowania dziur w obszarze *n* dzięki fluktuacjom termicznym i z dyfuzji dziur do obszaru *p*. Oba prądy, I_r i I_t, w warunkach równowagi termicznej są niewielkie i równe sobie, a zatem wypadkowy prąd przez złącze *p-n* w tych warunkach równa się zeru.

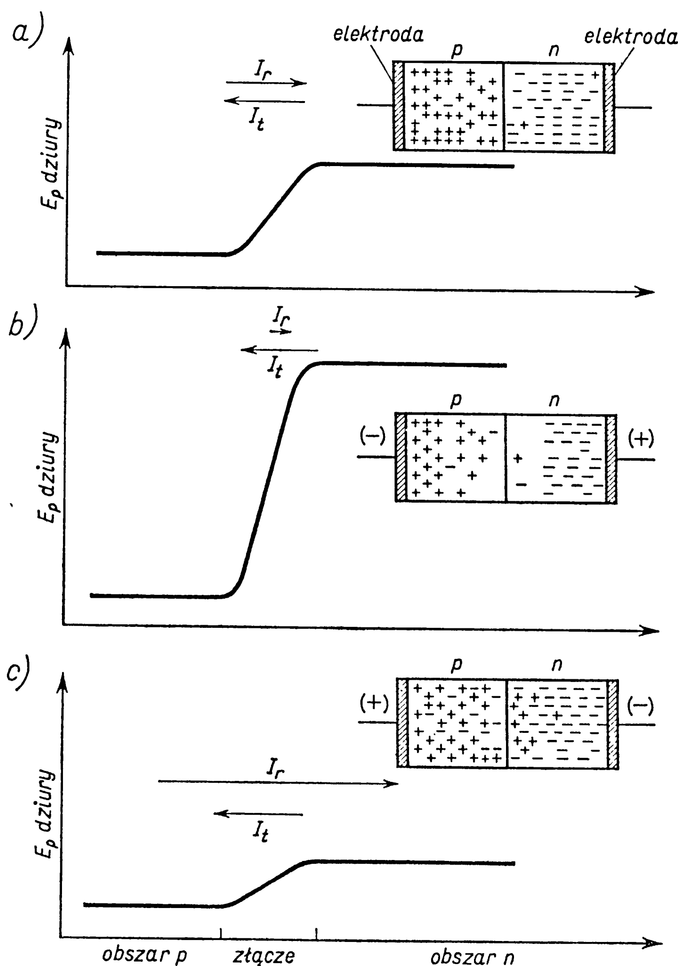
Omówioną wyżej sytuację na złączu *p-n* scharakteryzujemy z punktu widzenia modelu pasmowego. Pamiętamy, że do półprzewodników, podobnie jak do metali, zastosowaliśmy pojęcie poziomu Fermiego. W półprzewodnikach typu *p* był on bliski górnej granicy pasma walencyjnego, w półprzewodniku typu *n* znajdował się blisko dolnej granicy pasma przewodnictwa.

Z chwilą wytworzenia złącza *p-n*, na skutek omówionej wyżej dyfuzji nośników ładunku poprzez złącze następuje takie przesunięcie poziomów energetycznych półprze-

wodnika typu *p* i *n*, że ustala się wspólny dla obu typów poziom Fermiego. Równocześnie pojawia się na złączu między półprzewodnikami, w obszarze zwanym czasem warstwą zaporową, bariera potencjału i związana z nią różnica energii potencjalnej dziur (i elektronów) w obszarze typu *p* i *n* (rys. 32.18a).

32.16. Działanie prostujące złącza *p-n*

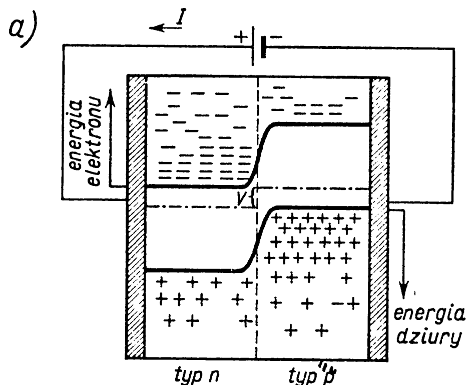
Działanie prostujące złącza *p-n*, podobnie jak działanie każdego innego urządzenia prostującego, sprowadza się do asymetrycznych zmian prądu elektrycznego przy zmianie znaku przyłożonego napięcia. Aby dokładniej zrozumieć tę asymetrię charakterystyki prądowo-napięciowej przy zastosowaniu złącza *p-n*, rozpatrzmy kolejno efekty towarzyszące przyłożeniu do złącza zewnętrznej różnicy potencjałów równej *V*, z tym

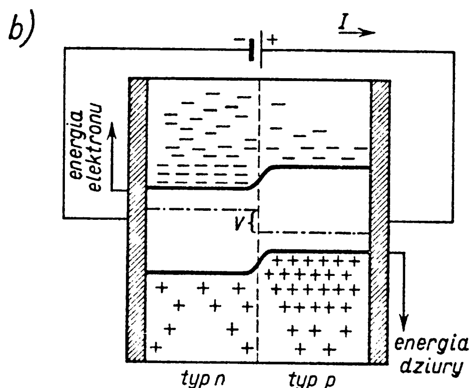


Rys. 32.18

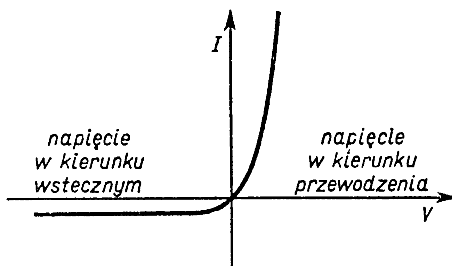
że w pierwszym przypadku dajemy napięcie V, jako tak zwane napięcie w kierunku wstecznym lub krótko napięcie wsteczne (+ na elektrodzie połączonej z półprzewodnikiem typu *n*), a w drugim — jako tak zwane napięcie w kierunku przewodzenia (+ na elektrodzie połączonej z półprzewodnikiem typu *p*).

Przyłożenie napięcia wstecznego (a więc + do obszaru *n* i — do obszaru *p*) odpowiada zwiększeniu „zaporowej" różnicy potencjałów między obu obszarami (rys. 32.18b). W tych warunkach prawdopodobieństwo przepływu nośników odpowiadającego prądowi rekombinacyjnemu (a więc np. dziur od *p* do *n*) spada prawie do zera: I_r staje się znikomo małe. Prąd termiczny natomiast pozostaje niezmieniony w stosunku do tego, który występował bez zewnętrznego pola. Przyłożenie napięcia w kierunku przewodzenia (a więc + do obszaru *p*, — do obszaru *n*) zmniejsza barierę potencjału zarówno dla dziur, jak i dla elektronów. Rośnie więc prąd I_r, podczas gdy prąd I_t pozostaje bez zmiany (32.18c). Rysunek 32.19a i b stanowi ilustrację opisanych zjawisk z punktu widzenia modelu pasmowego, a rys. 32.20 przedstawia typową charakterystykę prądowo-napięciową prostującego złącza *p-n*.





Rys. 32.19



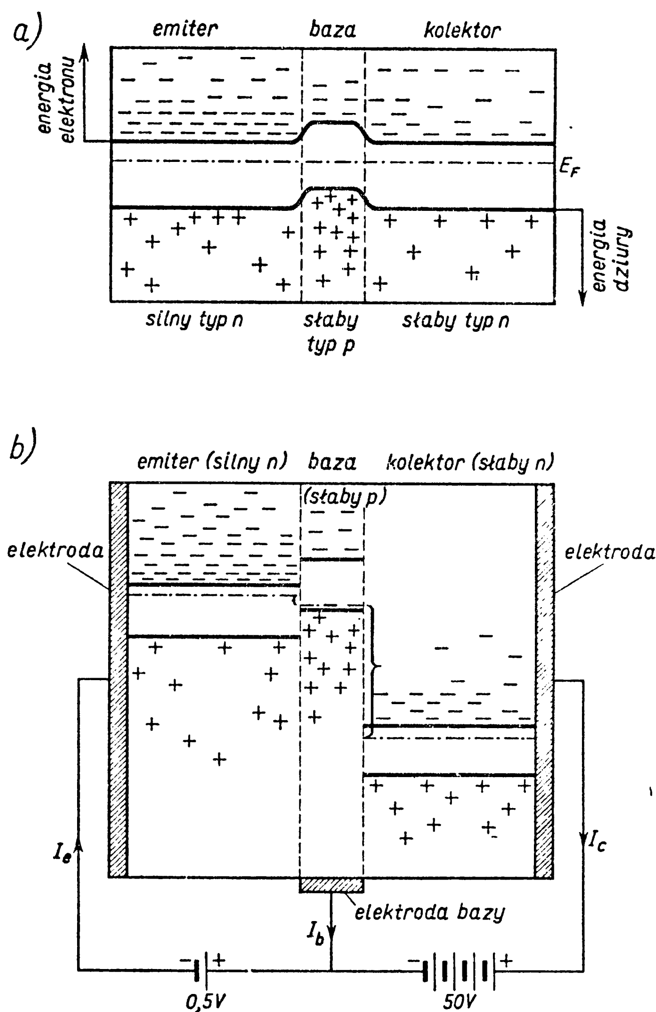
Rys. 32.20

Bardzo wyraźnie zaznaczona asymetria przebiegu charakterystyki prądowo-napięciowej świadczy o wysokich walorach prostujących opisanego urządzenia, nazywanego często diodą krystaliczną, gdyż spełnia tę samą funkcję co dwuelektrodowa lampa katodowa.

32.17. Tranzystor

Przejdziemy z kolei do omówienia zasady działania odpowiednika triody, zwanego *tranzystorem*, ograniczając się do przypadku tranzystora warstwowego *n-p-n* spełniającego zadanie wzmacniacza mocy.

Tranzystor tego typu przedstawiony jest schematycznie na rys. 32.21a i b. Rysunek a dotyczy tranzystora nie włączonego do obwodu elektrycznego, rysunek b — tranzystora włączonego do obwodu. Półprzewodnik środkowy o słabo zaznaczonym typie *p*, zwany *bazą*, połączony jest z lewej strony z „silnym" półprzewodnikiem typu *n*, zwanym *emiterem*, a z prawej strony — ze „słabym" półprzewodnikiem typu *n*, zwanym



Rys. 32.21

kolektorem. Na złączu między emiterem i bazą przyłożone jest stosunkowo niewielkie napięcie w kierunku przewodzenia o wartości rzędu kilku dziesiątych wolta. Jak już wiemy, w takich warunkach złącze stawia mały opór, prąd elektronowy osiąga dużą wartość.

Na złączu baza-kolektor przyłożone jest stosunkowo duże napięcie wsteczne, rzędu kilkudziesięciu woltów. Napięcie wsteczne decyduje o dużym oporze złącza i małym

prądzie elektronowym. Prąd ten w pojedynczej diodzie krystalicznej powinien być praktycznie niezależny od przyłożonego napięcia wstecznego (por. przebieg charakterystyki prądowo-napięciowej rys. 32.20).

Grubość obszaru bazy w tranzystorze jest bardzo mała, rzędu paru setnych milimetra. A zatem droga elektronu w obszarze bazy jest tak krótka, że małe jest prawdopodobieństwo rekombinacji elektronu z dziurą i w związku z tym prawie wszystkie elektrony wychodzące z emitera „przelatują" przez obszar bazy do kolektora, skąd już wobec bariery potencjału na złączu baza-kolektor nie mogą przedyfundować z powrotem do bazy. Dołączają się one do prądu elektronowego kolektora podnosząc znacznie jego wartość. Prąd płynący w obwodzie kolektora jest tylko nieco zmniejszony względem prądu z emitera. (Różnica obu prądów równa się małemu prądowi płynącemu z elektrody bazy.) Moc wyjściowa otrzymywana w obwodzie kolektora jest znacznie większa od mocy wejściowej w obwodzie emitera.

Przeanalizujmy nieco dokładniej zjawiska zachodzące w tranzystorze *n-p-n* włączonym do obwodu korzystając raz jeszcze z rys. 32.21b. Na rysunku tym zaznaczone są względne przesunięcia poziomów Fermiego: w emiterze i bazie o małe napięcie w obwodzie emitera, w bazie i kolektorze — o stosunkowo duże napięcie w obwodzie kolektora. Poza tym *zaznaczone są strzałkami następujące prądy elektronowe*: dopływający do elektrody emitera I_e, odpływający z elektrody bazy I_b oraz odpływający z elektrody kolektora I_c.

Elektrony dopływające od elektrody do emitera naruszają stan równowagi, charakterystyczny dla tranzystora, gdy brak jest zewnętrznego pola. Nadmiar elektronów w emiterze zwiększa zakres procesu rekombinacji elektronów z dziurami, ale w związku ze stosunkowo małą liczbą dziur w obszarze półprzewodnika donorowego nie prowadzi to do większej zmiany liczby aktywnych nośników prądu. Ważniejsze jest to, że w związku z wprowadzeniem nadmiaru elektronów wytwarza się gradient koncentracji elektronów powodujący dyfuzję elektronów do bazy.

Wobec włączenia na złączu emiter-baza napięcia w kierunku przewodzenia, elektrony mają nieco zmniejszoną barierę potencjału i łatwo wchodzą do obszaru bazy. Gdyby grubość bazy była duża, wchodzące do niej elektrony albo uległyby rekombinacji z dziurami, albo dotarłyby do elektrody bazy.

Oba efekty są niepożądane z punktu widzenia działania tranzystora, zależy nam bowiem na tym, aby możliwie nie zmniejszony prąd I_e dopłynął do kolektora.

Sprzyja temu — jak już mówiliśmy — bardzo mała grubość bazy, jak również to, że baza jest półprzewodnikiem o słabo zaakcentowanym typie *p*, a więc o stosunkowo małej koncentracji dziur. Tak więc większość elektronów z emitera dociera do kolektora, skąd tylko nieliczne pokonują barierę potencjału na złączu kolektor-baza i przyczyniają się do wytworzenia małego prądu bazy I_b.

Wprowadzamy następujące założenia upraszczające:

a) prąd bazy jest równy zeru, $I_b = 0$, wtedy $I_e = I_c$,

b) opór obwodu emiter-baza R_e jest stały i nieduży,

c) opór obwodu baza-kolektor R_c jest stały i duży,

d) opory pozostałych części obwodu elektrycznego można zaniedbać.

Przy tych założeniach zmianie prądu w obwodzie emitera o ΔI_e odpowiada zmiana mocy doprowadzającej równa $(\Delta I_e)^2 R_e$. Zmiana mocy występująca równocześnie w obwodzie kolektora równa się $(\Delta I_c)^2 R_c$. Wobec założenia, że $I_e = I_c$, moc wyjściowa rośnie w stosunku R_c/R_e. Zwiększenie mocy wyjściowej odbywa się tu oczywiście kosztem mocy doprowadzonej z baterii.

W ten sposób wyjaśniliśmy działanie tranzystora jako wzmacniacza. Zastosowania tranzystorów są szersze: stosuje się je w różnych generatorach, pulsatorach, układach przełącznikowych itp.

Tranzystory z powodzeniem zastępują w wielu urządzeniach lampy elektronowe. Wielką ich zaletą w stosunku do lamp elektronowych są ich małe wymiary, duża trwałość, wytrzymałość mechaniczna, mały pobór mocy.

Jedną z poważniejszych wad urządzeń półprzewodnikowych typu diod lub triod jest zależność ich działania od temperatury. Wiemy już, że do działania prostującego i wzmacniającego potrzebne są prostowniki typu *n* i *p* (połączone złączem), a więc półprzewodniki domieszkowe o nadmiarze elektronów lub dziur, decydującym o przesunięciu poziomu Fermiego ze środka przerwy energetycznej E_g w stronę pasma przewodnictwa lub pasma walencyjnego. Ale z drugiej strony wiemy też, że np. wzrost temperatury powoduje wzrost liczby elektronów w pasmie przewodnictwa i dziur — w pasmie walencyjnym. W pewnej temperaturze liczba termicznych nośników prądu może tak znacznie przewyższać liczbę nośników domieszkowych, że w zasadzie półprzewodnik zatraci swój pierwotny charakter domieszkowy, przekształcając się na półprzewodnik samoistny, dla którego poziom Fermiego znajdzie się w środku E_g. W tych warunkach przestanie istnieć złącze *p-n*, które decydowało o działaniu krystalicznej diody lub warstwowego tranzystora. To temperaturowe ograniczenie zakresu pracy omawianych urządzeń stanowi niewątpliwie ich dużą wadę.

Na zakończenie wróćmy raz jeszcze do różnic między przewodnikiem metalicznym, izolatorem i półprzewodnikiem. W przewodnikach metalicznych w każdej temperaturze mamy dużą liczbę nośników prądu w postaci elektronów. W izolatorach w temperaturach bliskich pokojowej nośników prądu jest bardzo mało. W półprzewodnikach dość duża liczba nośników występuje w temperaturach wysokich, mała natomiast w temperaturach niskich. Wzrost temperatury półprzewodnika wywołuje dwa procesy mające wpływ na zależność przewodności od temperatury. Jeden z nich to wzrost liczby nośników prowadzący do zwiększenia przewodności, drugi — to wzrost rozpraszania nośników na skutek oddziaływania ze strony sieci, prowadzący do zmniejszenia przewodności. W zależności od tego, który z tych efektów przeważa, w różnych obszarach temperaturowych w określonym półprzewodniku może występować wzrost lub spadek przewodności ze wzrostem temperatury. Częściej obserwowanym efektem ostatecznym jest wzrost przewodności półprzewodnika wraz ze wzrostem temperatury.

Pytania i zadania

1. Co jest charakterystyczne dla struktury elektronowej metali? Omówić strukturę elektronową kilku metali.

2. Jak się przedstawia model metalu oparty na koncepcji elektronów swobodnych? Omówić właściwości gazu elektronowego.

3. Na czym polega różnica między modelami klasycznym i kwantowym metali, opartymi na koncepcji elektronów swobodnych? Jakie dodatkowe fakty można było wyjaśnić po wprowadzeniu modelu kwantowego? Uzasadnić odpowiedź.

4. Jaka jest różnica w rozkładzie potencjału wewnątrz i na powierzchni metalu?

5. Wyjaśnić podstawy, na jakich opiera się model pasmowy metalu. Czym wywołana jest różnica w rozkładzie poziomów energetycznych izolowanych atomów i ciała stałego?

6. Jaka jest definicja pracy wyjścia? W jakich warunkach może wystąpić zjawisko emisji elektronu z powierzchni metalu? Scharakteryzować różne odmiany tego zjawiska i podać prawa nimi rządzące.

7. Co nazywamy potencjałem kontaktowym? Jak się zmieniają położenia poziomów Fermiego metali A i B w miejscu styku tych metali?

8. Co odróżnia metale, izolatory i półprzewodniki w ujęciu modelu pasmowego ciała stałego?

9. Czy istnieje ostra granica między izolatorami i półprzewodnikami? W jakich warunkach półprzewodnik jest izolatorem, a w jakich izolator może się stać półprzewodnikiem?

10. Jak się przedstawiają struktury: a) elektronowa, b) krystaliczna takich typowych półprzewodników jak krzem i german? Jak można scharakteryzować wiązania kowalencyjne?

11. Jaka jest rola domieszek z III i V grupy układu okresowego w germanie i krzemie?

12. Jak się zmienia przewodność półprzewodników w zależności od temperatury?

13. Scharakteryzować termistory i fotooporniki i podać ich najważniejsze zastosowania.

14. Uzasadnić nazwę: zjawisko fotoelektryczne wewnętrzne.

15. Czy selenek kadmu użyty jako fotoopornik w obwodzie elektrycznym wykaże zmiany oporu podczas zbliżania żarzącego się papierosa?

16. Omówić sens fizyczny pojęcia złącze *p-n*, zjawiska zachodzące na złączu, charakterystykę złącza z punktu widzenia modelu pasmowego: a) w warunkach równowagi termicznej bez zewnętrznego pola, b) z zewnętrznym polem.

17. Czy istnieje analogia między powstawaniem napięcia kontaktowego po złączeniu 2 metali i zjawiskami zachodzącymi na złączu *p-n* w równowadze termicznej, bez zewnętrznego pola elektrycznego. Przeprowadzić dokładną analizę obu zjawisk.

18. Omówić działanie prostujące złącza *p-n*. Podać charakterystykę prądowo-napięciową.

19. Omówić budowę i działanie tranzystora *n-p-n*.

20. Obliczyć gęstość swobodnych elektronów w miedzi w temperaturze 20°C, jeżeli wiadomo, że gęstość miedzi w tej temperaturze wynosi $8{,}89 \cdot 10^3$ kg/m^3, a masa atomowa 63,57.

Odp. $8{,}42 \cdot 10^{22}$.

21. Ile elektronów przewodnictwa zawiera 1 g sodu?

Odp. $2{,}6 \cdot 10^{22}$.

22. W zjawisku fotoelektrycznym wylatują z powierzchni cezu fotoelektrony o energii 2 eV. Jaka jest długość fali światła użytego do wywołania fotoemisji, jeżeli praca wyjścia dla cezu wynosi 1,8 eV.

Odp. 3240 Å.

23. Obliczyć prąd nasycenia płynący w lampie elektronowej, jeśli emiterem jest drucik wolframowy o długości 3 cm i średnicy 0,1 mm w temperaturze 2700K. Stała $A = 60{,}2$ A/cm$^2 \cdot$ K^2, praca wyjścia $W_w = 4{,}53$ eV.

Odp. około 0,147 A.

Uzupełnienia

U.1. Spójność układów jednostek

Spójność układów jednostek. Wzory wielkościowe i liczbowe. Dla zrozumienia istoty cechy spójności układu jednostek rozważymy zapis pomiaru dowolnej wielkości fizycznej oznaczonej np. symbolem x i wprowadzimy rozróżnienie równań wielkościowych i liczbowych.

Jak wiadomo, dowolny pomiar polega na porównaniu badanej wielkości z jednostką miary danej wielkości. Liczba otrzymana z tego porównania stanowi liczbową wartość miary danej wielkości. Symbolicznie oznaczymy ją $\{x\}$. Zmierzona wielkość jest jednoznacznie określona (innymi słowy, znana jest miara wielkości), gdy podana jest wartość liczbowa i jednostka miary. Jednostkę miary wielkości x oznaczamy $[x]$, ewentualnie z zaznaczeniem układu jednostek, np. $[x]_{SI}$.

A zatem zapisem miary wielkości x w określonym układzie jednostek „a" jest

$$x = \{x\}[x]_a.$$

Pojęcie miary wielkości x można traktować jako iloczyn liczby $\{x\}$ i jednostki miary $[x]$. Oczywiście zapis miary wielkości zmienia się w zależności od wyboru jednostki: im większa jest jednostka, tym mniejsza jest wartość liczbowa miary danej wielkości. Innymi słowy, obowiązuje proporcja

$$\frac{\{x\}_a}{\{x\}_b} = \frac{[x]_b}{[x]_a}.$$

Równanie to może oddawać usługi przy przeliczaniu wartości liczbowych z jednych układów na inne, a więc np. z układów jednostek dotychczas stosowanych na układ SI.

Zajmiemy się obecnie zapisem równań fizycznych. Przypuśćmy, że w wyniku badań ilościowych ustalona została następująca zależność

$$y = f(x_1; x_2; x_3; \dots x_n). \tag{1}$$

Równanie to wyraża, że wielkość pochodna y jest funkcją wielkości $x_1, x_2, x_3, \dots x_n$.

Jeżeli każdy z występujących w tym wzorze symboli będziemy traktować jako miarę danej wielkości (tzn. jako iloczyn wartości liczbowej i jednostki miary), to wypisany wzór będzie tzw. *wzorem wielkościowym*. Można go napisać w postaci rozwiniętej:

$$\{y\}[y] = f(\{x_1\}[x_1]; \{x_2\}[x_2]; \{x_3\}[x_3]; \dots \{x_n\}[x_n]). \tag{2}$$

Jeżeli natomiast każdy z symboli we wzorze (1) będzie przedstawiać tylko wartość liczbową danej wielkości, to ten wzór będzie *wzorem liczbowym*. Można go napisać w postaci rozwiniętej:

$$\{y\} = kf(\{x_1\}; \{x_2\}; \{x_3\}; \dots \{x_n\}), \tag{3}$$

gdzie k oznacza współczynnik liczbowy zależny od stosowanych jednostek.

Ponieważ takie rozszerzone zapisy wzorów jak (2) i (3) są rzadko stosowane, należy w zasadzie zawsze wyraźnie podkreślać, czy zapis wzoru należy traktować jako wielkościowy, czy też jako liczbowy. Wzory fizyczne podane w tej książce mają charakter wzorów wielkościowych.

Wyobraźmy sobie, że tworzymy pewien układ jednostek oparty na kilku jednostkach podstawowych. Jeśli równanie (2) potraktujemy jako równanie definicyjne nowej wielkości y, to w wyborze jednostki tej nowej wielkości będziemy mieli pewną swobodę. Jeżeli jednak narzucimy warunek, że jednostka ta ma być zdefiniowana w ten sposób, by współczynnik k w równaniu (3) był równy jednostce, to równanie liczbowe przyjmie oczywiście najprostszą postać. Jeżeli taki warunek narzucimy wszystkim kolejno definiowanym jednostkom wielkości pochodnych, to układ jednostek złożony z przyjętych jednostek podstawowych i z opisanych jednostek pochodnych będzie stanowił układ zwany spójnym. Tę niesłychanie ważną cechę, upraszczającą postać równań liczbowych, ma układ SI (podobnie zresztą jak układ CGS i MKS). Należy jednak pamiętać, że cechy spójności nie miałby już układ, w którym stosowane byłyby różne wielokrotne lub podwielokrotne jednostki SI.

Wiadomości dotyczące wzorów wielkościowych i liczbowych oraz spójności układów jednostek zastosujemy do dowolnego prostego wzoru fizycznego, Na przykład prędkość v ciała w ruchu jednostajnym można przedstawić jako

$$v = \frac{s}{t}.$$

Traktując to wyrażenie jako wzór wielkościowy napiszemy (w postaci rozwiniętej):

$$\{v\}[v] = \frac{\{s\}[s]}{\{t\}[t]}. \tag{4}$$

Zależność ta jest słuszna zawsze, niezależnie od wyboru jednostek miary $[v]$, $[s]$ i $[t]$.

Po prostym przekształceniu (4) otrzymujemy

$$\{v\} = \frac{[s]\{s\}}{[v][t]\{t\}}$$

lub

$$v = k\frac{\{s\}}{\{t\}}, \tag{5}$$

gdzie

$$k = \frac{[s]}{[v][t]}.$$

Wzór (5) jest wzorem liczbowym. Występujący w nim współczynnik może przyjmować różne wartości, zależnie od wyboru jednostek miary $[v]$, $[s]$ i $[t]$. Będzie on równy

1, jeśli wszystkie wielkości będą wyrażone w jednostkach należących do spójnego układu jednostek. Przy zastosowaniu jednostek niespójnych współczynnik k może przyjmować bardzo różne wartości. Jeśli np. $[s]$ i $[t]$ należą do układu CGS, a $[v]$ do układu SI, to

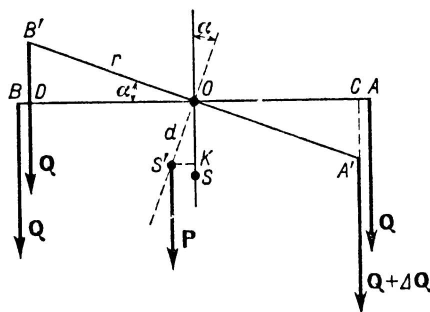
$$k = \frac{1}{\mathrm{m/s}} \cdot \frac{\mathrm{cm}}{\mathrm{s}} = \frac{1}{100}.$$

Współczynnik k występujący we wzorach liczbowych i zależny od stosowanych jednostek miary należy odróżniać od współczynników liczbowych występujących w wielu wzorach empirycznych oraz od współczynników liczbowych związanych z definicjami pewnych wielkości.

U.2. Ważenie

W punkcie niniejszym omówimy pojęcie czułości wagi i podamy wskazówki praktyczne dotyczące dokładnego ważenia.

U.2.1. *Czułość wagi.* Załóżmy, że w czasie ważenia dzięki odpowiedniemu doborowi odważników osiągnęliśmy stan równowagi wagi. W tych warunkach dołożenie małego odważnika ΔQ powoduje wychylenie belki od położenia pierwotnego o mały kąt α, zależny od rodzaju wagi. Stosunek $\alpha/\Delta Q$ (lub w przybliżeniu $\operatorname{tg}\alpha/\Delta Q$) nazywamy czułością wagi. *Czułość wagi* wyraża liczbowo kąt wychylenia belki przypadający na jednostkowe małe obciążenie jednej z szalek, odpowiednio dobrane do rodzaju wagi. Na przykład w przypadku wagi analitycznej może to być wychylenie belki przypadające na 1 mg obciążenia. W przypadku wagi technicznej może to być wychylenie przypadające na 1 dg itp.



Rys. U.2.1

Zastosowaniu nadwyżki obciążenia ΔQ towarzyszy przesunięcie środka ciężkości belki S (rys. U.2.1) do punktu S'. Dzięki temu powstaje dodatkowy moment obrotowy siły ciężkości i belka pozostaje w równowadze w nowym odchylonym od pierwotnego położeniu.

Celem wyprowadzenia wzoru na czułość wagi wypisujemy równanie określające

równość momentów sił działających na oba ramiona wagi oznaczając długość ramion wagi przez r oraz odległość OS' środka ciężkości od osi podparcia wagi przez d:

$$Q\cdot OD+P\cdot S'K = (Q+\Delta Q)OC,$$

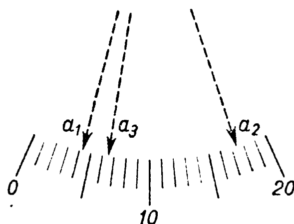
$$Qr\cos\alpha+Pd\sin\alpha = (Q+\Delta Q)r\cos\alpha,$$

$$dP\sin\alpha = \Delta Qr\cos\alpha,$$

skąd

$$\frac{\operatorname{tg}\alpha}{\Delta Q} = \frac{r}{Pd}. \tag{U.2.1}$$

Wyprowadzenie wzoru na czułość wagi było proste dzięki temu, że przyjęliśmy, iż punkty A, O i B leżą na jednej prostej (czyli, że ostrza podparcia belki i zawieszeń szalek leżą w jednej płaszczyźnie). Gdyby punkt podparcia belki był przesunięty, to otrzymalibyśmy inny wzór na czułość wagi. Wtedy czułość zależałaby nie tylko od długości ramion r, ciężaru belki i szalek P oraz odległości d środka ciężkości od punktu podparcia, lecz także od Q, czyli od ciężaru ciała ważonego. Nowoczesne wagi mają budowę taką, że A, O i B leżą na jednej prostej (dokładniej mówiąc, osi obrotu szalek i belki leżą w jednej płaszczyźnie). Czułość takich wag, jak widać z wzoru U.2.1, jest tym większa, im dłuższe są ramiona, im mniejszy jest ciężar belki i szalek, im bliżej osi podparcia znajduje się środek ciężkości. Czułości wagi nie można dowolnie zwiększać przez uwzględnienie tych zależności. Ciężar belki można zmniejszyć do pewnych granic przez nadanie jej postaci kraty, ale wzrost długości ramion pociąga za sobą wzrost ciężaru belki i wzrost okresu wahań. Zbliżenie środka ciężkości do punktu podparcia ma wpływ na stałość równowagi, zwiększa okres wahań, utrudnia i przedłuża proces ważenia.



Rys. U.2.2

U.2.2. *Dokładne ważenie.* Ustalanie na oko położenia poziomego belki jako położenia równowagi nie prowadziłoby do dokładnych wyników ważenia. Postępujemy więc inaczej, wykorzystujemy mianowicie dodatkową wskazówkę, zwaną czasem języczkiem wagi. Wskazówka ta prostopadła do belki jest przymocowana w jej środku (prostopadle do niej). Podczas wahań belki koniec wskazówki przesuwa się na tle podziałki. Aby ustalić położenie wskazówki odpowiadające równowadze wagi nieobciążonej (zwane miejscem zerowym), unosimy belkę za pomocą aretażu tak, aby mogła wykonywać wahania. Przy niewielkich wahaniach belki notujemy np. trzy kolejne wychylenia a_1, a_2 i a_3 języczka wagi na tle podziałki (rys. U.2.2). Uwzględniamy nieparzystą liczbę wychyleń języczka w związku z zanikającym charakterem wahań wskutek

nieuniknionego tarcia. W naszych rozważaniach ograniczymy się do dwóch wychyleń a_1 i a_3 w jedną stronę i jednego a_2, pośredniego, w drugą stronę. Z wielkości tych trzech wychyleń znajdujemy miejsce zerowe A_0 wagi nieobciążonej

$$A_0 = \frac{\frac{a_1+a_3}{2}+a_2}{2}.$$

Zważenie jakiegoś ciała sprowadza się do takiego dobrania odważników, aby miejsce zerowe A wagi obciążonej pokryło się z miejscem zerowym A_0 wagi nieobciążonej. Jest to łatwe w przypadku wag mało czułych: w przypadku wag dokładniejszych, bardziej czułych, trwałoby to zbyt długo. Stosujemy wtedy inną metodę. Dokładny wynik ważenia znajdujemy stosując tzw. interpolację.

Ważenie z interpolacją polega na tym, że masę (ciężar) ciała określamy raz z nadmiarem, drugi raz — z niedomiarem.

Niech masie m^+ wyznaczonej z nadmiarem odpowiada miejsce zerowe A^+, a masie m^- wyznaczonej z niedomiarem odpowiada miejsce zerowe A^-. Miejsca zerowe A^+ i A^- wyznaczamy podobnie jak A_0. Fakt, że jedno ważenie jest z nadmiarem, a drugie z niedomiarem poznajemy po tym, że wartość A_0 jest zawarta między A^- i A^+. Jeśli ciało ważone umieszczone jest na szalce lewej i pierwszą podziałkę od lewej strony uważamy za zerową (por. rys. U.2.2), to

$$A^+ < A_0 < A^-.$$

Jeśli wyjątkowo wypadnie $A_0 = A^+$ (lub A^-), to m^+ (lub m^-) jest dokładną masą m ciała ważonego i wtedy przeprowadzenie interpolacji jest niepotrzebne.

Masy m^+ i m^- powinny się różnić o najmniejszy odważnik używany przy danym rodzaju wagi, a więc np. o 0,1 g w przypadku wagi technicznej, o 0,01 g w przypadku wagi analitycznej (bez konika).

Przeprowadźmy na przykład rozważania interpolacyjne dla wagi technicznej. Zakładamy, że przesunięcia miejsc zerowych są proporcjonalne do różnicy obciążeń szalek, a zatem jeśli różnica obciążeń $m^+ - m^- = 0{,}1$ g wywołuje zmianę miejsca zerowego o $A^- - A^+$, to różnica obciążeń $m^+ - m = x$ wywołałaby zmianę miejsca zerowego o $A_0 - A^+$, skąd

$$x = 0{,}1\left|\frac{A_0-A^+}{A^- - A^+}\right|.$$

Szukaną masę m ciała ważonego znajdziemy odejmując znalezioną wartość x od m^+:

$$m = m^+ - x, \quad \text{czyli} \quad m = m^+ - 0{,}1\left|\frac{A_0-A^+}{A^- - A^+}\right|.$$

Analogiczne rozważania w przypadku ważenia z niedomiarem prowadzą do wzoru:

$$m = m^- + 0{,}1\left|\frac{A^- - A_0}{A^- - A^+}\right|.$$

Wynik ostateczny na m jest oczywiście niezależny od tego czy rozważania dotyczą ważenia z nadmiarem, czy też z niedomiarem.

Obliczenia interpolacyjne wykonujemy tak, aby otrzymać wynik z dokładnością o jedno miejsce dziesiętne większą niż daje bezpośrednie wyznaczenie masy m^+ lub m^- za pomocą odważników. Tak więc przez interpolację na wadze technicznej otrzymujemy dokładność do 0,01 g, na wadze analitycznej (bez konika) do 0,001 g. Tę dokładność uzyskaną na wadze analitycznej możemy powiększyć przez zastosowanie tzw. konika.

Konik jest to cienki drucik skręcony tak, że może być wygodnie zawieszony w różnych punktach na belce. Masa konika jest dokładnie równa 10 mg. Konik umieszczony na szalce, ewentualnie zawieszony na końcu belki działa jak zwykły odważnik dziesięciomiligramowy. Ramiona belki wagi analitycznej podzielone są na 10 równych części. Zawieszenie konika na $\frac{1}{10}$ długości ramienia, licząc od punktu podparcia, czyli na pierwszej kresce od środka belki, powoduje powstanie momentu siły 10 razy mniejszego niż wtedy, gdy konik spoczywa na szalce. Konik zawieszony na pierwszej kresce tak się przyczynia do wychylenia belki, jak odważnik 1-miligramowy umieszczony na szalce. Umieszczenie konika na $\frac{2}{10}$ długości ramienia (na drugiej kresce licząc od środka belki) wywołuje moment siły równy $\frac{2}{10}$ momentu odpowiadającego położeniu konika na szalce. Tym razem działanie konika jest takie, jak działanie odważnika 2-miligramowego położonego na szalce.

Zaopatrzenie wagi analitycznej w konika powiększa jak gdyby komplet odważników stwarzając możliwości stosowania obciążeń 1 mG, 2 mG, 3 mG, ... itd., odpowiednio do numeru porządkowego podziałki, na której zawieszony jest konik.

Zastosowanie interpolacji podczas ważenia na wadze analitycznej z konikiem wymaga określenia miejsca zerowego wagi nieobciążonej oraz miejsc zerowych przy obciążeniu z nadmiarem i niedomiarem. Te dwa obciążenia powinny się różnić o 1 mG, tzn. przechodząc od jednego obciążenia do drugiego przesuwamy konik o jedną podziałkę na belce. Podczas ważenia z interpolacją obowiązuje, jak zwykle, sprawdzenie, czy A_0 jest zawarte między A^+ i A^-.

Stosowanie interpolacji podczas ważenia na wadze analitycznej z konikiem daje dokładność do 0,0001 G (dokładność wyznaczania masy do 0,0001 g).

U.3. Lampy elektronowe

U.3.1. *Uwagi ogólne.* Przystępując do omówienia zasad budowy i działania lamp elektronowych zestawimy na początku pewne cechy wspólne dla różnych typów, a dopiero w dalszym ciągu dodamy nieco szczegółów charakterystycznych dla poszczególnych typów i ich zastosowań.

Lampy elektronowe są to bańki próżniowe o wysokiej próżni. Ciśnienie jest w nich mniejsze od 10^{-3}–10^{-4} N/m^2 (czyli od 10^{-5}–10^{-6} Tr; 1 Tr = 1 mm Hg). W każdej lampie występują dwie podstawowe elektrody: katoda i anoda. Katoda ma najczęściej postać drucika z niklu lub jego stopów, pokrytego tlenkami baru i strontu. Anoda (zwykle z blachy żelaznej lub niklowej) ma kształt cylindra otaczającego katodę. Katodę łączymy ze źródłem prądu niskiego napięcia i doprowadzamy do żarzenia. Materiał katody jest tak dobrany, że praca wyjścia elektronów jest stosunkowo niewielka. Dzięki

temu już przy niezbyt wysokich temperaturach otrzymujemy stosunkowo dużą emisję elektronów.

W niektórych lampach elektronowych mamy inny typ katody, a mianowicie katodę żarzoną pośrednio, w postaci pustego cylindra z niklu z dodatkiem baru i strontu. Wewnątrz cylindra znajduje się odizolowany element grzejny, np. w postaci spirali wolframowej żarzonej prądem. Spirala ogrzewa cylinder, który wysyła elektrony. Katody żarzone pośrednio mają tę zaletę, że cała ich powierzchnia jest powierzchnią ekwipotencjalną.

Bardzo ważnym czynnikiem warunkującym dobrą pracę lamp jest utrzymywanie w lampie niskiego ciśnienia. Odpompowywanie powietrza z baniek odbywa się na gorąco, aby ułatwić wyjście gazów zaadsorbowanych na szkle i metalu. Mimo to w czasie pracy lampy resztki gazów zaadsorbowanych mogłyby się wydzielać i powodować wzrost ciśnienia. Aby temu zapobiec stosuje się specjalne pochłaniacze (tzw. gettery), zawierające magnez lub bor, które wprowadza się do bańki i po odpompowaniu powietrza silnie ogrzewa. Dzięki temu metal paruje, a para osiada na ściance bańki w postaci cienkiej warstwy ciemnego nalotu. Warstwa ta silnie adsorbuje wydzielające się gazy i pomaga w utrzymaniu wysokiej próżni.

U.3.2. *Diody*. Rysunek U.3.1 odnosi się do najprostszej lampy dwuelektrodowej i przedstawia jej najprostszy obwód z uwzględnieniem jej symbolicznego oznaczenia. W środku bańki mamy drucik K stanowiący katodę, żarzony z baterii B_1 oraz anodę A połączoną z dodatnim biegunem źródła wysokiego napięcia B_2.

Żarzony drucik K otoczony jest przestrzennym ładunkiem ujemnym wynikającym z termoemisji. Wobec istnienia pola elektrycznego między anodą i katodą odbywa się przepływ elektronów do anody. Prąd w znaczeniu umownym płynie od potencjału wyższego do niższego, a więc od anody do katody. Natężenie tego prądu zależy od napięcia między A i K zwanego napięciem anodowym oraz od temperatury katody. Przy stałej temperaturze katody w miarę wzrostu napięcia anodowego natężenie prądu stopniowo rośnie, coraz więcej elektronów w jednostce czasu odpływa z ładunku przestrzennego w otoczeniu katody do anody. W końcu jednak natężenie prądu osiąga wartość maksymalną — natężenie prądu nasycenia. Zależność natężenia prądu od napięcia przedstawiają krzywe na rys. U.3.2a i b. Na rysunku b zaznaczone są trzy krzywe dla różnych temperatur katody w postaci drucika wolframowego. Jak widać, wyższym temperaturom odpowiadają większe natężenia prądu nasycenia. Przebieg krzywych zależny jest od materiału, z którego zrobiona jest katoda (praca wyjścia). Wykresy przedstawiające zależność natężenia prądu anodowego od napięcia anodowego noszą nazwę *charakterystyki lampy dwuelektrodowej*. Badania wykazują, że dla małych napięć między anodą i katodą obowiązuje zależność

$$I_a = kU_a^{3/2}, \qquad \text{(U.3.1)}$$

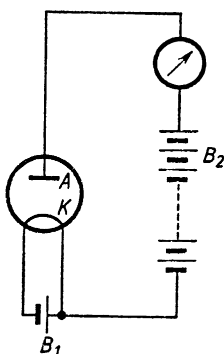
przy czym współczynnik k zależy od rozmiarów i kształtu elektrod, a nie zależy od temperatury katody. Dlatego też początkowa część krzywych charakterystyk ma wspólny przebieg, niezależnie od temperatury katody.

Wartość pochodnej dI_a/dU_a nosi nazwę *nachylenia charakterystyki lampy*. Jest ona

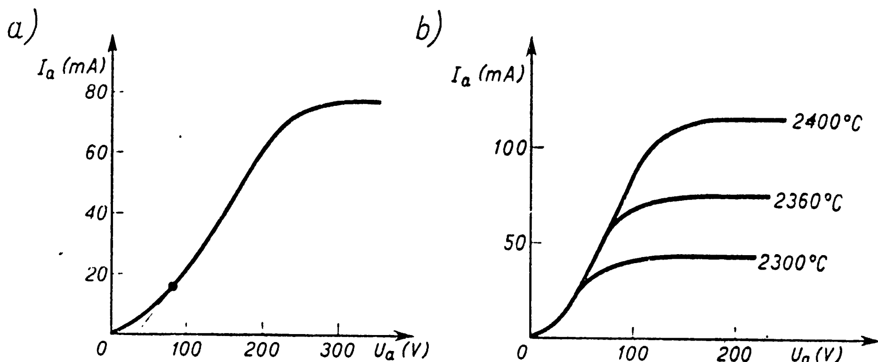
równa tangensowi kąta, który tworzy styczna do charakterystyki z osią odciętych. Nachylenie charakterystyki w przypadku małych napięć otrzymujemy różniczkując wyrażenie (U.3.1):

$$\frac{dI_a}{dU_a} = \tfrac{3}{2} k \sqrt{U_a}.$$

Jak widać z przebiegu charakterystyk na rys. U.3.2, ich nachylenie początkowo zmienia się dość szybko, następnie przy potencjałach większych można charakterystykę uważać za prostoliniową, a jej nachylenie za stałe.



Rys. U.3.1

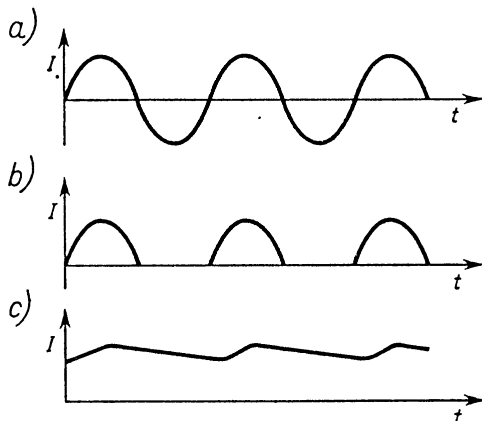


Rys. U.3.2

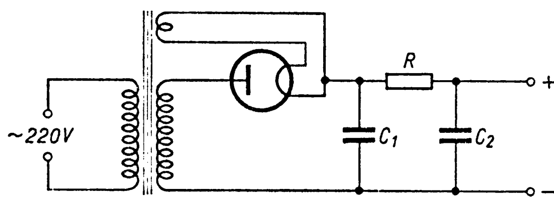
Lampom dwuelektrodowym przypisujemy pewien opór wewnętrzny. Gdyby przebieg charakterystyki był od początku prostoliniowy, to opór określalibyśmy zgodnie z prawem Ohma i traktowalibyśmy jako stały. Wobec nieprostoliniowego przebiegu charakterystyki opór wewnętrzny lampy definiujemy jako

$$R_w = \frac{dU_a}{dI_a}. \qquad \text{(U.3.2)}$$

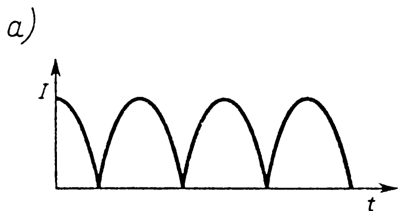
Działanie prostownicze diody (czyli lampy dwuelektrodowej) wyjaśnić można następująco. Załóżmy, że do zimnej elektrody przyłożyliśmy niższy potencjał niż do elektrody rozżarzonej (innymi słowy, zmieniliśmy role anody i katody). W tych warunkach pole elektryczne będzie przyciągało elektrony do powierzchni żarzonego drutu i prąd elektryczny przez lampę nie będzie płynął. A zatem po włączeniu między anodę i katodę źródła prądu zmiennego otrzymamy w obwodzie tylko prąd jednokierunkowy, czyli zamiast przebiegów sinusoidalnych prądu (rys. U.3.3a) powstaną przebiegi przedstawione na rys. U.3.3b. Będzie to tzw. jednokierunkowe prostowanie prądu.

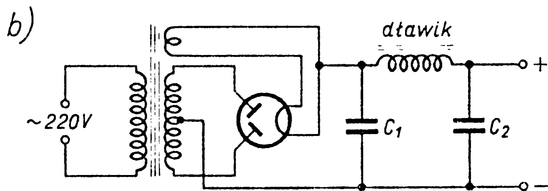


Rys. U.3.3



Rys. U.3.4





Rys. U.3.5

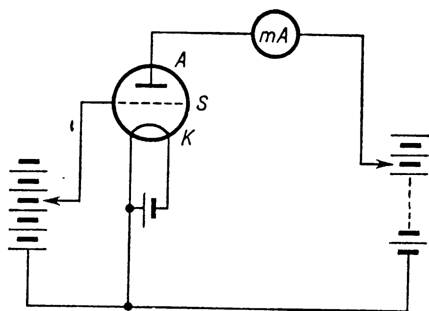
Schemat układu elektrycznego do jednokierunkowego prostowania przedstawiony jest na rys. U.3.4. Włączone dodatkowo pojemności i opór zmniejszają w pewnym stopniu wahania napięcia otrzymywanego prądu (rys. U.3.3c).

W celu otrzymywania prądu jednokierunkowego, o charakterze takim jak na rys. U.3.5a, stosuje się obwody (rys. U.3.5b) o dwukierunkowym prostowaniu, których

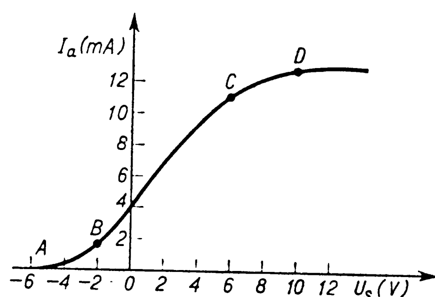
zasadniczą część stanowi lampa elektronowa o dwóch anodach. Podczas jednej połowy okresu prądu zmiennego prąd płynie od jednej anody do katody, podczas następnej połowy okresu — od anody drugiej do katody. Zastosowane dodatkowo kondensatory i dławik lub opór doprowadzają do tzw. wygładzenia prądu, w którym wahania napięcia są minimalne.

U.3.3. *Triody*. Znacznie szersze zastosowanie niż dioda mają lampy trójelektrodowe, zwane triodami. Schemat budowy triody widoczny jest na rysunku U.3.6a. Różnica między diodą i triodą polega na wprowadzeniu do lampy trzeciej elektrody tzw. siatki S, najczęściej w postaci metalowej spirali otaczającej katodę. Rysunek U.3.6 przedstawia układ lampy trójelektrodowej (służący do wykreślania charakterystyki triody), z uwzględnieniem jej symbolicznego oznaczenia.

W triodzie natężenie prądu płynącego od anody do katody zależy nie tylko od napięcia anodowego (anody względem katody), lecz także — i to w dużym stopniu — od napięcia siatkowego (siatki względem katody). Wpływ siatki na ruch elektronów jest bardzo duży w związku z jej małą odległością od katody. Jeżeli siatka ma potencjał ujemny względem katody, to odpycha elektrony. Gdy ujemny potencjał siatkowy jest dostatecznie duży, ruch elektronów może być całkowicie zahamowany. Mówimy wtedy o zablokowaniu lub zatkaniu lampy. Gdy siatka ma potencjał dodatni, to przyciąga dodatkowo elektrony wzmacniając w ten sposób działanie anody. Większość elektronów przebiega między zwojami spirali i dochodzi do anody, czyli dodatni potencjał siatki wpływa na zwiększenie prądu anodowego.



Rys. U.3.6



Rys. U.3.7

Zmieniając napięcie siatkowe U_s w obwodzie przedstawionym na rys. U.3.6 i mierząc odpowiadające mu natężenie I_a prądu anodowego, możemy wykreślić charakterystykę statyczną triody. Charakterystyka statyczna triodv, czyli krzywa wyrażająca zależność I_a od U_s:

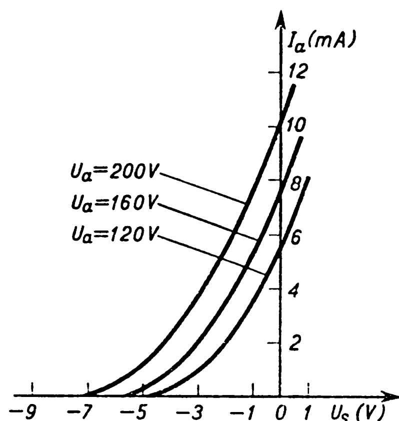
$$I_a = f(U_s), \quad \text{gdy} \quad U_a = \text{const},$$

przedstawiona jest na rys. U.3.7. Przy wykreślaniu charakterystyki statycznej napięcie anodowe pozostaje niezmienne.

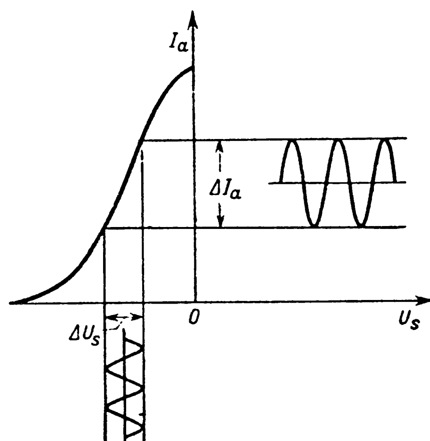
Jak widać z rys. U.3.7, na krzywej charakterystyki można wyróżnić część krzywoliniową AB, następnie część o przebiegu prawie prostoliniowym BC, której odpowiada

prosta proporcjonalność natężenia prądu anodowego do napięcia siatkowego, dalej po lekkim zakrzywieniu na odcinku *CD* część odpowiadającą natężeniu nasycenia. W tej ostatniej części charakterystyki wzrostowi napięcia siatkowego nie towarzyszy wzrost natężenia prądu anodowego, gdyż wszystkie elektrony wylatujące z rozżarzonej katody dochodzą już do anody.

Po zmianie napięcia anodowego otrzymuje się krzywą (U.3.8) o analogicznym charakterze, lecz przesuniętą w lewo w przypadku zwiększenia napięcia anodowego, a w prawo — w przypadku jego zmniejszenia.



Rys. U.3.8



Rys. U.3.9

Zaletą triody w stosunku do diody jest możliwość silniejszego wpływania na prąd anodowy przez zmianę napięcia siatkowego niż przez analogiczne zmiany napięcia anodowego w diodzie. W zastosowaniach triody korzystamy z tego, że bardzo mały ładunek doprowadzony do siatki, zmieniający jej potencjał zaledwie o parę woltów, wywołuje dość znaczne zmiany prądu anodowego, rzędu kilku mA. Wobec stosunkowo wysokich napięć anodowych wiąże się to z dużą zmianą mocy. Na tym polega zastosowanie triody jako wzmacniacza np. w odbiornikach radiowych do wzmacniania słabych prądów powstających w antenie. Wpływ zmian napięcia siatkowego na natężenie prądu anodowego widoczny jest na rys. U.3.9. Załóżmy, że załączamy napięcie siatkowe U_s o stosunkowo małych wahaniach ΔU_s (np. z anteny, mikrofonu, adaptera), przypadające np. w obszarze prostoliniowej części charakterystyki triody. Wykres charakterystyki uzupełniamy dwiema dodatkowymi częściami, a mianowicie rysujemy dwie osie czasu, poziomą i pionową, jako osie symetrii wahań ΔI_a i ΔU_s. Następnie wykreślamy krzywe wahań napięcia i natężenia prądu. Wahania natężenia prądu przy danych wahaniach napięcia są tym większe, im bardziej stromo przebiega krzywa charakterystyki. Nachylenie prostoliniowej części charakterystyki triody określa w pewnym sensie jej „czułość", tzn. stopień reakcji na wahania napięć siatkowych. *Nachylenie S prostoliniowej części*

charakterystyki definiujemy jako tangens kąta utworzonego przez odcinek prostoliniowy charakterystyki z osią odciętych. A zatem

$$S = \left(\frac{\Delta I_a}{\Delta U_s}\right)_{U_A} = \text{const.} \tag{U.3.3}$$

Innymi słowy, w celu wyznaczenia nachylenia charakterystyki obliczamy stosunek odpowiadających sobie przyrostów ΔI_a i ΔU_s przy niezmiennym napięciu anodowym U_a. Jak widać ze wzoru, nachylenie może być podawane w mA/V.

Drugą wielkością charakteryzującą triodę jest jej opór wewnętrzny. *Opór wewnętrzny lampy* definiujemy jako

$$R = \left(\frac{\Delta U_a}{\Delta I_a}\right)_{U_s} = \text{const} \tag{U.3.4}$$

czyli jako stosunek odpowiadających sobie zmian napięcia i natężenia anodowego przy niezmiennym potencjale siatki.

Trzecią wielkością charakteryzującą lampę jest *współczynnik wzmocnienia* (*amplifikacji*), który oznaczamy przez K. Zwiększając potencjał siatki o ΔU_s powodujemy wzrost natężenia prądu anodowego o ΔI_a. Tę samą zmianę natężenia prądu można by wywołać zmieniając napięcie anodowe o pewną wartość ΔU_a, lub — inaczej rozumując — można powiedzieć, że w celu zachowania stałego natężenia prądu anodowego należałoby zmniejszyć napięcie anodowe o ΔU_a. Właśnie stosunek

$$K = -\left(\frac{\Delta U_a}{\Delta U_s}\right)_{I_a} = \text{const}, \tag{U.3.5}$$

jest współczynnikiem wzmocnienia lampy.

Odwrotność współczynnika wzmocnienia lampy nazywamy *przechwytem*:

$$D = \frac{1}{K} = -\left(\frac{\Delta U_s}{\Delta U_a}\right)_{I_a} = \text{const.}$$

Przechwyt D wyrażamy zwykle w procentach.

Łatwo sprawdzić, że dla określonej lampy iloczyn nachylenia S, oporu wewnętrznego R i przechwytu D, czyli $SRD = 1$.

W triodach współczynnik wzmocnienia waha się w granicach 10–60 i zależy głównie od odległości siatki od katody, poza tym od gęstości zwojów spirali siatki (liczby zwojów na jednostce długości) i grubości drutu spirali. Opór wewnętrzny waha się od 10 000–50 000 Ω. Prąd nasycenia wynosi 10–40 mA.

W przypadku lampy o dużej mocy (20 kW) mamy inne dane charakterystyczne: $K = 40$, $S = 25$ mA/V, $R \approx 1600\ \Omega$, prąd nasycenia ≈ 15 A, robocze napięcie anodowe $\approx 10\,000$ V. „Nić" katody stanowi drut wolframowy o średnicy 1 mm. Prąd żarzenia osiąga wartość 50 A.

Lampa trójelektrodowa wykorzystywana jest w układzie wzmacniacza oporowego. Obwód jednolampowego wzmacniacza oporowego przedstawiony jest na rys. U.3.10. Prócz znanych nam już elementów do obwodu anodowego włączony jest duży opór

R_a na ogół znacznie większy od oporu wewnętrznego R_w lampy. Zmienne napięcie U_1 przeznaczone do wzmocnienia załączamy do obwodu siatki, a z zacisków oporu R_a pobieramy napięcie wzmocnione U_2.

W tych warunkach napięcie anodowe U_a można wyrazić jako różnicę siły elektromotorycznej E_a baterii anodowej i spadku potencjału na oporze anodowym $I_a R_a$, czyli

$$U_a = E_a - I_a R_a.$$

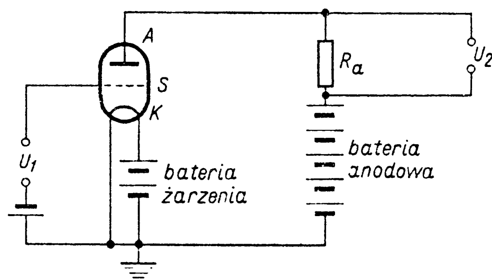
Jak widać z tego równania, gdy wzrost potencjału siatki wywoła zwiększenie prądu anodowego, to automatycznie zmniejszy się potencjał anody. Można udowodnić, że

$$-\left(\frac{\Delta U_a}{\Delta U_s}\right) = \frac{K}{1+\frac{R_w}{R_a}}.$$

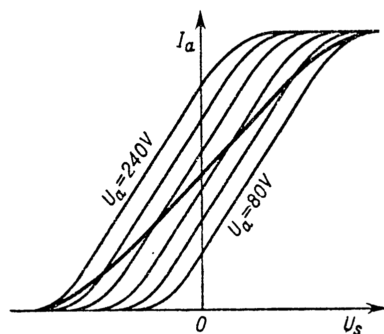
Gdy R_a jest znacznie większe od R_w, to

$$K \approx -\left(\frac{\Delta U_a}{\Delta U_s}\right).$$

Znak minus przypomina, że wzrostowi potencjału siatki odpowiada zmniejszanie się potencjału anody i odwrotnie.



Rys. U.3.10



Rys. U.3.11

Jeśli opór włączony do obwodu anodowego obejmuje nie tylko rezystancję, lecz także reaktancję pojemnościową i indukcyjną, to zamiast R_a wstawiamy moduł impedancji:

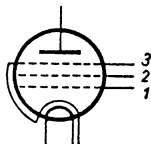
$$Z = \sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}.$$

Z wzoru wyrażającego moduł impedancji widać, że wzmocnienie będzie w tych warunkach zależne od pulsacji ω prądu doprowadzonego do siatki.

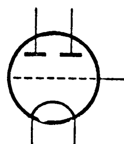
Warto zwrócić uwagę na jeszcze jedną sprawę, a mianowicie, gdy w obwodzie anody znajduje się opór, to podczas pracy lampy — jak to już podkreślano — zmianom ΔU_s towarzyszy zmiana napięcia anodowego. Wykreślając zależność $I_a = f(U_s)$ przy stałym oporze R_a i stałej sile elektromotorycznej E_a baterii anodowej otrzymamy nową charakterystykę lampy, zwaną dynamiczną (rys. U.3.11 — linia grubo nakreślona na

tle rodziny charakterystyk statycznych). Nachylenie tej nowej charakterystyki będzie mniejsze od nachylenia S.

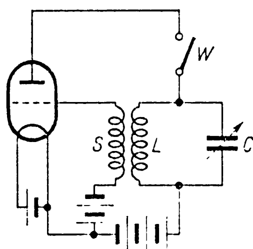
U.3.4. *Lampy wieloelektrodowe.* Ustaliliśmy, że zmiany U_s prowadzą do zmian I_a, a zmiany I_a wywołują wahania U_a. Duże wahania potencjału anody mogą jednak indukcyjnie (elektrostatycznie) oddziaływać na siatkę, a tym samym zakłócać pracę lampy. Aby usunąć te niekorzystne oddziaływania, anodę osłania się drugą siatką II o stałym potencjale (mniej więcej $\frac{1}{2}U_a$). Druga *siatka*, zwana *ekranującą*, spełnia więc rolę osłony elektrostatycznej. Siatka I zachowuje swą poprzednią rolę i nazywa się *siatką sterującą*. W ten sposób powstaje lampa zwana *tetrodą*.



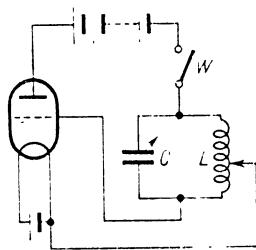
Rys. U.3.12


Rys. U.3.13

W związku z zastosowaniem siatki II występuje jednak nowe, niekorzystne zjawisko. Szybko pędzące elektrony dolatując do powierzchni anody wybijają z jej powierzchni nowe elektrony (czyli wywołują tzw. emisję wtórną), które biegną do dodatniej siatki II, a więc zmniejszają prąd anodowy w obwodzie. Aby temu zapobiec, wprowadza się jeszcze trzecią *siatkę*, tzw. *chwytną*, umieszczoną między anodą a siatką ekranującą i połączoną z katodą wewnątrz lampy lub poza nią. Lampa o trzech siatkach, anodzie i katodzie, czyli o pięciu elektrodach, nosi nazwę *pentody* (rys. U.3.12). Pentody są obecnie bardzo szeroko stosowane. W radiotechnice znalazły zastosowanie lampy o jeszcze większej liczbie elektrod, jak również lampy podwójne, które w jednej bańce zawierają dwa układy lamp, np. diodę i triodę, dwie triody itp. (rys. U.3.13).



Rys. U.3.14



Rys. U.3.15

U.3.5. *Lampa elektronowa jako generator drgań niegasnących.* Lampa elektronowa użyta jako generator drgań niegasnących musi być sprzężona indukcyjnie (rys. U.3.14) lub bezpośrednio (rys. U.3.15) z układem elektrycznym drgającym o pojemności C i indukcyjności L. Rozważmy przykładowo obwód triody o sprzężeniu indukcyjnym. Po zamknięciu klucza W rozpoczyna się stopniowe ładowanie kondensatora C, aż do osiągnięcia napięcia rozładowania między okładkami. W czasie rozładowania w obwo-

dzie LC powstają drgania gasnące, które nakładają się na stały prąd anodowy. Aby zapobiec ich wygaśnięciu, umieszczamy w pobliżu cewki L cewkę S włączoną do obwodu siatkowego lampy. Przy takim zestawieniu cewka L działa indukcyjnie na cewkę S wytwarzając zmienny potencjał siatkowy, który wywołuje synchroniczne wahania prądu anodowego. Dzięki odpowiedniemu zestawieniu obwodu prąd anodowy podsyca drgania zachodzące w układzie LC. Zastosowanie kondensatora o zmiennej pojemności pozwala na zmianę częstotliwości własnej układu drgającego i częstotliwości wytwarzanej przez generator drgań.

Warunkiem otrzymania silnych drgań w układzie drgającym LC jest uzyskanie rezonansu układu drgającego i układu lampy, musi zatem zachodzić równość

$$LC = L_1 C_1 .$$

W związku z tym w układzie drgającym uwzględniono kondensator o zmiennej pojemności.

Chcąc odbierać drgania wytworzone przez generator należy zastosować nowy układ drgający, będący rezonatorem drgań, a więc taki, dla którego znowu spełniony będzie warunek rezonansu. W układzie rezonatora musi również występować kondensator o zmiennej pojemności, aby można było pulsację drgań własnych tego układu doprowadzić do wartości

$$\omega = \sqrt{\frac{1}{LC}} .$$

Tabela U.1*

Wielkości mechaniczne, ich symbole i jednostki

L.p.	Wielkość	Symbol wielkości	Jednostka miary w układzie SI		Jednostki dopuszczone przejściowo do stosowania (wybrane)		Relacje między jednostkami
			nazwa	oznaczenie	nazwa	oznaczenie	
1	2	3	4	5	6	7	8
1	długość	l	metr	m	mikron	μ	10^{-6} m
	szerokość	b			angstrem	Å	10^{-10}m
	wysokość	h					
	grubość	d					
	promień	r					
	średnica	d					
	długość drogi	s					
	długość łuku	s					
2	pole powierzchni	S	metr kwadratowy	m^2			
3	objętość	V	metr sześcienny	m^3			
	pojemność						
4	czas	t, τ	sekunda	s	minuta	min	60 s
					godzina	h	3600 s
					doba	d	86400 s
5	kąt płaski	α, β...	radian	rad	stopień kątowy	...°	0,017453rad
					obrót (kąt pełny)	obr	6,283185 rad
6	kąt bryłowy	Ω, ω	steradian	sr			
7	prędkość liniowa	v	metr na sekundę	m/s	metr na minutę	m/min	1/60 m/s
	prędkość światła	c			kilometr na godzinę	km/h	1/3,6 m/s
8	przyspieszenie liniowe	a	metr na kwadrat sekundy	m/s^2			
	przyspieszenie ziemskie	g					
9	prędkość kątowa	ω	radian na sekundę	rad/s	obrót na sekundę	obr/s	2π rad/s
10	przyspieszenie kątowe	α	radian na kwadrat sekundy	rad/s^2	obrót na kwadrat sekundy	obr/s^2	2π rad/s^2
11	okres	T	sekunda	s			
12	częstotliwość	ν	herc	Hz			
13	masa	m	kilogram	kg	tona	t	1000 kg
14	gęstość (masa właściwa)	ϱ	kilogram na metr sześcienny	kg/m^3	gram na centymetr sześcienny	g/cm^3	1000 kg/m^3

* Tablice U.1–U.4 obejmują tylko wielkości i jednostki miary wymieniane w podręczniku.

Tabela U. 1 (c.d.)

1	2	3	4	5	6	7	8
15	pęd (ilość ruchu)	p	kilogramometr na sekundę	kg · m/s			
16	moment pędu (kręt)	L	kilogram razy metr kwadratowy na sekundę	$kg \cdot m^2/s$			
17	moment bezwładności	I	kilogram razy metr kwadratowy	$kg \cdot m^2$			
18	siła	F,P,Q	niuton	N	dyna	dyn	$10^{-5}N$
	ciężar	P			kilogram-siła (kilopond)	kG, kp	9,80665 N
	tarcie	T					
19	ciężar właściwy	γ	niuton na metr sześcienny	N/m^3			
20	moment siły	M	niutonometr	N · m	kilogram-siła razy metr	kG · m	9,80665 N · m
	moment kierujący	D					
21	ciśnienie	p	niuton na metr kwadratowy	N/m^2	bar	bar	$10^5 N/m^2$
	naprężenie normalne	σ			tor (milimetr słupa rtęci)	Tr (mmHg)	$133,322\ N/m^2$
	naprężenie styczne	τ			atmosfera techniczna	at	$98066,5\ N/m^2$
					atmosfera normalna (fizyczna)	atm	$101325\ N/m^2$
					kilogram-siła na metr kwadratowy	kG/m^2	$9,80665\ N/m^2$
					kilogram-siła na centymetr kwadratowy	kG/cm^2	$98066,5\ N/m^2$
22	moduły sprężystości	E,G, K	jak w p. 21				
23	lepkość dynamiczna	η	niutonosekunda na metr kwadratowy	$N \cdot s/m^2$	puaz	P	$0,1\ N \cdot s/m^2$

Tabela U.1 (c.d.)

1	2	3	4	5	6	7	8
24	napięcie powierzchniowe	σ	niuton na metr	N/m	dyna na centymetr	dyn/cm	0,001 N/m
25	praca	W	dżul	J	erg	erg	10^{-7}J
	energia	E, Q			kilogramometr	kGm	9,80665 J
	energia potencjalna	E_p			kilowatogodzina	kWh	$3{,}6 \cdot 10^{6}$J
	energia kinetyczna	E_k			kaloria	cal	4,1868 J
					elektronowolt	eV	$1{,}602 \cdot 10^{-19}$J
26	moc	P	wat	W	erg na sekundę	erg/s	10^{-7} W
	strumień energii	Φ			kilogramometr na sekundę	kGm/s	9,80665 W
					koń mechaniczny	KM	735,4987 W
27	gęstość strumienia energii	φ	wat na metr kwadratowy	W/m²			
28	natężenie pola grawitacyjnego	g	niuton na kilogram	N/kg			
29	potencjał pola grawitacyjnego	V	dżul na kilogram	J/kg			
30	amplituda drgań	A, A_0 B	metr	m			
31	stała tłumienia	δ	odwrotność sekundy	s^{-1}			
32	współczynnik oporu ośrodka	b	kilogram na sekundę	kg/s			
33	logarytmiczny dekrement tłumienia	Λ					
34	długość fali	λ	metr				
35	współczynnik tarcia:						
	statycznego	μ_0					
	dynamicznego (kinematycznego)	μ					
36	względny przyrost:						
	długości	ε					
	objętości	θ					
37	skręcenie proste	γ					
38	liczba Poissona	μ					
39	współczynnik dyfuzji	D	metr kwadratowy na sekundę	m²/s			

Tabela U.2

Wielkości cieplne, ich symbole i jednostki

L.p.	Wielkość	Symbol wielkości	Jednostka miary w układzie SI		Jednostki dopuszczone przejściowo do stosowania (wybrane)		Relacja między jednostkami
			nazwa	oznaczenie	nazwa	oznaczenie	
1	2	3	4	5	6	7	8
1	temperatura termodynamiczna	T	kelwin	K			$T = t + 273,15$
					stopień Celsjusza	°C	$\Delta T = \Delta t$*
2	współczynnik rozszerzalności: liniowej objętościowej	α γ	odwrotność kelwina	K^{-1}			
3	ilość ciepła ciepło energia wewnętrzna	Q U	dżul	J	kaloria	cal	4,1868 J
4	współczynnik przewodzenia ciepła	λ	wat na metr i na kelwin	W/m · K	kaloria na sekundę, na centymetr i na kelwin	$\frac{cal}{s \cdot cm \cdot K}$	418,68 W/m · K
5	pojemność cieplna	C	dżul na kelwin	J/K	kaloria na kelwin	cal/K	4,1868 J/K
6	ciepło właściwe ciepło właściwe pod stałym ciśnieniem ciepło właściwe w stałej objętości	c c_p c_V	dżul na kilogram i na kelwin	J/kg · K	kaloria na gram i na kelwin	cal/g · K	418,6 J/kg · K
7	ciepło topnienia ciepło parowania	c_t r	dżul na kilogram	J/kg	kaloria na gram	cal/g	4186,8 J/kg
8	entropia	S	dżul na kelwin	J/K	kaloria na kelwin	cal/K	4,1868 J/K
9	współczynnik prężności	β	odwrotność kelwina	K^{-1}			
10	sprawność	η					

* Do wyrażania różnic temperatur należy stosować oznaczenie K.

Tabela U.2 (c.d.)

1	2	3	4	5	6	7	8
11	prędkość najprawdopodobniejsza	$\hat{\vartheta}$	metr na sekundę	m/s			
12	średnia droga swobodna	l	metr	m			
13	stopień swobody	f					

Tabela U.3

Wielkości elektryczne i magnetyczne, ich symbole i jednostki

L.p.	Wielkość	Symbol wielkości	Jednostka miary w układzie SI		Jednostki dopuszczone przejściowo do stosowania (wybrane)		Relacje między jednostkami
			nazwa	oznaczenie	nazwa	oznaczenie	
1	2	3	4	5	6	7	8
1	prąd elektryczny	I	amper	A			
2	gęstość prądu	j	amper na metr kwadratowy	A/m^2	amper na milimetr kwadratowy	A/mm^2	$10^6 A/m^2$
3	ładunek elektryczny	Q	kulomb	C	amperogodzina	A · h	3600 C
4	gęstość powierzchniowa ładunku	σ	kulomb na metr kwadratowy	C/m^2			
5	natężenie pola elektrycznego	E	wolt na metr niuton na kulomb	V/m N/C			1N/C = 1V/m
6	potencjał elektryczny napięcie elektryczne siła elektromotoryczna	V U E	wolt	V			
7	indukcja elektryczna	D	kulomb na metr kwadratowy	C/m^2			
8	strumień indukcji elektrycznej strumień elektryczny	φ_D φ_E	kulomb niuton razy metr kwadratowy na kulomb	C $N \cdot m^2/C$			

Tabela U.3 (c.d.)

1	2	3	4	5	6	7	8
9	pojemność elektryczna	C	farad	F			
10	przenikalność elektryczna (stała dielektryczna)	ε	farad na metr	F/m			1F/m = = 1C²/N · m²
	przenikalność elektryczna próżni	ε_0					
	przenikalność elektryczna względna	ε_r	—				
11	moment dipolowy elektryczny	p	kulombometr	C · m			
12	natężenie pola magnetycznego	H	amper na metr	A/m	ersted	Oe	$\frac{1}{4}\cdot 10^3$A/m
	koercja	H_k					
13	indukcja magnetyczna	B	tesla	T	gaus	Gs	10^{-4}T 1T = 1Wb/m²
	pozostałość magnetyczna	B_0					
14	strumień magnetyczny	Φ	weber	Wb	makswel	Mx	10^{-8}Wb 1Wb = 1V · s
15	indukcyjność własna	L	henr	H			1H = 1Ω · s
	indukcyjność wzajemna	L_{12}					
16	przenikalność magnetyczna	μ	henr na metr	H/m			
	przenikalność magnetyczna próżni	μ_0					
	przenikalność magnetyczna względna	μ_r	—				
17	opór elektryczny rezystancja	R	om				
	impedancja (moduł)	Z					
18	moment elektromagnetyczny	μ	amper razy metr kwadratowy	A · m²			

Tabela U.4

Wielkości optyczne, ich symbole i jednostki

L.p.	Wielkość	Symbol wielkości	Jednostka miary w układzie SI: nazwa	Jednostka miary w układzie SI: oznaczenie	Relacje między jednostkami
1	strumień energii promienistej	Φ_e	wat	W	
2	gęstość kątowa strumienia energii promieniowania	I_e	wat na steradian	W/sr	
3	natężenie napromieniowania	E_e	wat na metr kwadratowy	W/m^2	
4	luminancja energetyczna	L_e	wat na metr kwadratowy i steradian	$W/m^2 \cdot sr$	
5	światłość kierunkowa (gęstość kątowa strumienia energii promieniowania widzialnego)	I	kandela	cd	
6	strumień świetlny	Φ	lumen	lm	$1\ lm = 1\ cd \cdot sr$
7	natężenie oświetlenia	E	luks	lx	$1\ lx = 1\ \frac{lm}{m^2}$
8	luminancja	L	nit	nt	$1\ nt = 1\ \frac{cd}{m^2}$
9	współczynnik załamania	n			
10	ogniskowa	f	metr	m	
11	powiększenie liniowe	W_l			
12	powiększenie kątowe	W_k			
13	długość fali	λ	metr	m	
14	kąt łamiący pryzmatu	φ			
15	kąt odchylenia	δ			
16	liniowy współczynnik absorpcji	μ	odwrotność metra	m^{-1}	
17	zdolność emisji	e	wat na metr kwadratowy	W/m^2	
18	zdolność absorpcji	α			
19	zdolność odbijania	ϱ			
20	zdolność przepuszczania	τ			

Tabela U.5

Podstawowe stałe fizyczne

Wielkość	Oznaczenie i wartość w układzie SI
stała grawitacyjna	$G = 6{,}67 \cdot 10^{-11}$ N · m²/kg²
przyspieszenie ziemskie normalne	$g_n = 9{,}80\,665$ m/s²
stała gazowa	$R = 8{,}3143$ J/mol · K
stała Boltzmanna	$k = 1{,}380\,54 \cdot 10^{-23}$ J/K
prędkość światła w próżni	$c = 2{,}997\,925 \cdot 10^{8}$ m/s
stała Stefana–Boltzmanna	$\sigma = 5{,}6697 \cdot 10^{-8}$ W/m² · K⁴
stała Wiena	$= 2{,}8978 \cdot 10^{-3}$ m · K
przenikalność elektryczna próżni	$\varepsilon_0 = 8{,}854\,15 \cdot 10^{-12}$ C²/N · m²
przenikalność magnetyczna próżni	$\mu_0 = 4\pi \cdot 10^{-7}$ Wb/A · m
stała Faradaya	$F = 9{,}648\,70 \cdot 10^{4}$ C/mol
stała Avogadra	$N_A = 6{,}022\,52 \cdot 10^{23}$ 1/mol
ładunek elementarny	$e = 1{,}602\,10 \cdot 10^{-19}$ C
masa elektronu	$m_e = 9{,}109\,1 \cdot 10^{-31}$ kg
masa neutronu	$m_n = 1{,}674\,82 \cdot 10^{-27}$ kg
masa protonu	$m_p = 1{,}672\,52 \cdot 10^{-27}$ kg
promień elektronu	$r_e = 2{,}817\,77 \cdot 10^{-15}$ m
stała Plancka	$h = 6{,}625\,6 \cdot 10^{-34}$ J · s
	$\hbar = 1{,}054\,50 \cdot 10^{-34}$ J · s
stała Rydberga	$R_\infty = 1{,}097\,373 \cdot 10^{7}$ m⁻¹
	$R_H = 1{,}096\,775 \cdot 10^{7}$ m⁻¹
magneton Bohra	$\mu_B = 9{,}273\,2 \cdot 10^{-24}$ J/T
moment magnetyczny elektronu	$\mu_e = 9{,}284\,0 \cdot 10^{-24}$ J/T

Skorowidz nazwisk

Skorowidz rzeczowy